

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 11:32:54 ; Search time 29.74 Seconds
(without alignments)
1000.934 Million cell updates/sec

Title: US-09-674-235-1

Sequence: 1406
1 MAFERLSQVGLGRFQMLH.....RKISRKRHKNDYKVKTKF 268

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*

1: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
21: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*
22: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1406	100.0	268	21	AAV52386	Human transmembran
2	1238	88.1	541	22	AAAB47276	HOAT5. Homo sapie
3	1237	88.0	391	22	ABBI1732	Human secreted pro
4	993	70.6	213	22	AAW24188	Human EST encoded
5	799.5	56.9	553	22	AAE04897	Human transporter
6	768	54.6	554	22	AAAB47275	HOAT4. Homo sapie
7	765	54.4	341	22	AAW40059	Human polypeptide
8	759	54.0	339	21	AAAB19406	Amino acid sequenc
9	601	42.7	553	22	AAE10332	Human transporter
10	594.5	42.3	369	22	AAW41845	Human polypeptide
11	592	42.1	550	22	AAE06612	Human protein havi

12	592	42.1	550	22	AAAB69091	Human organic anio
13	590.5	42.0	566	22	AAE10336	Human transporter
14	588.5	41.9	578	22	AAE06571	Human protein havi
15	472	33.6	645	22	AAU32379	Novel human secret
16	458.5	32.6	550	21	AAI44278	Human organic anio
17	458.5	32.6	550	22	AAAB47271	HOAT1. Homo sapie
18	458.5	32.6	553	20	AAW88489	Human organic anio
19	451.5	32.1	551	20	AAW88488	Rat organic anion
20	436.5	31.0	545	22	AAAB36553	Mouse organic anio
21	428	30.4	815	22	ABG26899	Novel human diagno
22	390.5	27.8	542	22	AAAY92902	HOAT3. Homo sapie
23	386.5	27.5	542	22	AAAB47274	Novel human secret
24	373	26.5	128	22	AAU33066	Mouse osteoclast t
25	367.5	26.1	537	18	AAW44195	Rat cerebral organ
26	362.5	25.8	536	21	AAAY92903	Peptide #11836 enc
27	357.5	25.4	561	18	AAW44196	Peptide #11836 enc
28	344	24.5	113	22	ABAB44330	Protein #9189 enco
29	344	24.5	113	22	ABAB27190	Human brain expres
30	344	24.5	113	22	AAW5378	Human bone marrow
31	344	24.5	113	22	AAW8056	Peptide #8388 enco
32	344	24.5	113	22	AAW21954	Peptide #12329 enc
33	344	24.5	113	22	AAW38292	Rat liver anion tr
34	324.5	23.1	535	21	AAV51249	Murine organic an
35	308	21.9	540	22	AAAB9401	HOAT2B. Homo sapi
36	307.5	21.9	538	22	AAAB47273	HOAT2A. Homo sapi
37	307.5	21.9	546	22	AAAB47272	A human organic an
38	296.5	21.1	546	21	AAAB08823	Human liver specif
39	296.5	21.1	552	22	ABBI12030	A human organic an
40	295	21.0	551	21	AAAB08824	Novel human diagno
41	289.5	20.6	358	22	AAW36897	Human reproductive
42	286.5	20.4	139	22	AAAB12131	Hydrophobic domain
43	267	19.0	607	21	AAAB12131	Novel human diagno
44	231.5	16.5	264	22	ABG27003	Drosophila melano
45	228	16.2	538	22	ABAB60174	

ALIGNMENTS

RESULT 1
ID AAV52386 standard; Protein: 268 AA.
XX AAV52386;
AC
XX
XX 09-FEB-2000 (first entry)
XX
XX Human transmembrane protein HP02000.
DE
XX
XX HP02000; transmembrane domain; liver; expression; homology;
KW organic cation transporter; drug excretion; antibody; assay reagent;
KW diagnostic marker; primer; probe; antisense; gene therapy;
KW agonist; antagonist; ligand; therapeutic.
XX
XX Homo sapiens.
OS
XX
XX WO9955862-A2.
PN
XX
XX 04-NOV-1999.
PD
XX
XX 27-APR-1999; 99WO-JP02226.
PF
XX
XX 28-APR-1998; 98JP-0119395.
PR
XX
XX (SAGA) SAGAMI CHEM RES CENT.
PA (PROT-) PROTEGENE INC.
XX
XX Kato S, Kimura T;
PI
XX WPI: 2000-023358/02.
XX N-PDB: AA238316, AA238317.
DR
XX
XX Human proteins with transmembrane domains, involved in control of cell

PT proliferation and differentiation, useful for treating e.g. cancer or
 PT inflammation

PS Claim 1: Page 74-75; 11app; English.

XX This sequence represents the protein encoded by human cDNA clone
 CC HP02000 which is a 32 kD in size with two putative transmembrane
 CC domains. The cDNA was isolated from a human liver cell cDNA library, and
 CC from tissue localisation studies has been found to be expressed only in
 CC the liver. The protein has homology with the rat organic cation
 CC transporter (EMBL Accession No. Y09945) which is involved in drug
 CC excretion, and this sequence may have a similar function. The protein
 CC may be used to raise specific antibodies, as assay reagents, as
 CC ligands and binding proteins, for the isolation of cognate receptors,
 CC Nucleotides encoding the protein may be used as primers and probes or
 CC antisense molecules, and in gene therapy. Cells transformed with these
 CC nucleotides may be used to screen for agonists and antagonists which are
 CC potentially useful therapeutically.

XX Sequence 268 AA:

Query Match 100.0%; Score 1406; DB 21; Length 268;
 Best Local Similarity 100.0%; Pred. No. 1.2e-143;
 Matches 268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MAFELLISQVGLGRFOMLHVFLIPSLMLIPHILENFAAIPGRCWVHMLDNTGS 60
 Db 1 mafeellisqvglgrfomlvflipsmliphillenfaaipghrcwvhlmdntgs 60
 OY 61 GNEGTLSIEDALLRISIPDSNLRPEKCRFPVHPOMQLHNGTHTSSEADTEPCVDGW 120
 Db 61 gnegtlsiedallrisipdsnlrpekcrrfvpwqllhngtlhtsseatdepcvdgw 120
 OY 121 VYDOSYFPTITVKMDLYCDYQSLKSVQFLLTGMLVGGIIGHVSDFRWVESARWLI 180
 Db 121 vydosyfpstivtkmdlycdyqslksvqfllltgmlvvgiigghvsdfrwvesarwli 180
 OY 181 TNKDEGKALKRKARVINGIKNAEETLNIEVVRSTMOEELDAQRTTYCDLFRNDSMRK 240
 Db 181 tnkdegkalkrkavringiknaeetlnievvrstmoeeldaqrttycdlfrnpsmrk 240
 OY 241 RICILVLRKIRSRHRNDCTYKTFK 268
 Db 241 ricilvlrkkisrrhrndctykvckf 268

RESULT 2

AAB47276
 ID AAB47276 standard; Protein: 541 AA.

XX AAB47276;

DT 06-AUG-2001 (first entry)

XX hoAT5.

XX Human; organic anion transporter; hoAT; liver; kidney;
 KW membrane protein; transport; organic anion; splice variant.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 282 /note= "PKC phosphorylation site"
 FT Modified-site 289 /note= "PKC phosphorylation site"
 FT Modified-site 345 /note= "PKC phosphorylation site"
 FT Modified-site 526 /note= "PKC phosphorylation site"
 FT /note= "PKC phosphorylation site"

PN WO200104283-A2.

PD 18-JAN-2001.

PF 12-JUL-2000; 2000WO-US18980.

PR 12-JUL-1999; 99US-0143771.

PA (META-) METABASIS THERAPEUTICS INC.

XX Sun W;

XX WPI: 2001-367057/38.

DR N-PSDB: AAC85824.

FT Nucleic acids encoding human organic anion transporter polypeptides,
 PT useful in gene therapy procedures -

XX Claim 2; Fig 6; 95pp; English.

XX The sequences given in AAB47271-76 represent human organic anion
 CC transporter (hoAT) polypeptides. hoAT polypeptides are preferentially
 CC expressed in the liver and kidneys of humans. OAT's are membrane
 CC proteins that facilitate the transport of organic anions across the
 CC cell membrane. The mechanism of transport is thought to be a secondary
 CC cell membrane. The mechanism of transport is thought to be a secondary
 CC hoAT2 and hoAT3 are thought to be splice variants as they are
 CC identical except at the C-terminal end. hoAT proteins and the DNA
 CC encoding them, may be used in the prevention, treatment and diagnosis
 CC of diseases associated with inappropriate hoAT expression.

XX Sequence 541 AA:

Query Match 88.1%; Score 1238; DB 22; Length 541;
 Best Local Similarity 69.7%; Pred. No. 5.2e-125;
 Matches 249; Conservative 0; Mismatches 0; Indels 108; Gaps 1;

OY 1 MAFELLISQVGLGRFOMLHVFLIPSLMLIPHILENFAAIPGRCWVHMLDNTGS 60
 Db 1 mafeellisqvglgrfomlvflipsmliphillenfaaipghrcwvhlmdntgs 60
 OY 61 GNEGTLSIEDALLRISIPDSNLRPEKCRFPVHPOMQLHNGTHTSSEADTEPCVDGW 120
 Db 61 gnegtlsiedallrisipdsnlrpekcrrfvpwqllhngtlhtsseatdepcvdgw 120
 OY 121 VYDOSYFPTITVKMDLYCDYQSLKSVQFLLTGMLVGGIIGHVSDFRWVESARWLI 168
 Db 121 vydosyfpstivtkmdlycdyqslksvqfllltgmlvvgiigghvsdfrwvesarwli 168
 OY 169 -----
 Db 169 -----
 OY 181 IqlaIdcaafptlfvycvllflagfssmltsnslptewlrpskallvllssga 240
 Db 169 -----RMVESARWLIITNKLDGKALR 192
 Db 241 lslglllgglayvfrdcwqlhvvasvflgllllgtrwvesarwllitnkldgkalkr 300
 OY 193 KYARTGINKNAEETLNIEVVRSTMOEELDAQRTTYCDLFRNDSMRKRICIIVFLR 249
 Db 301 kvartngiknaeetlnievvrstmoeeldaqrttycdlfrnpsmrkricilivflr 357

RESULT 3

AAB11732
 ID AAB11732 standard; peptide; 391 AA.

XX AAB11732;

DT 11-JAN-2002 (first entry)

XX Human secreted protein homologue, SEQ ID NO:2102.

KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW hematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; hematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antistatic; antiatheritic; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnereary; antitumor.
 XX Homo sapiens.
 OS
 XX WO200157188-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 05-FEB-2001; 2001WO-US03800.
 PE
 XX 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Drmanac RT;
 PI
 XX WPI: 2001-457740/49.
 DR N-PSDB; ABA08976.
 XX
 XX Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 e.g. arthritis and cancer -
 PT
 XX
 XX Claim 20; Page 235-236; 1963pp; English.
 PS
 XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug

CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.
 CC
 XX
 SQ Sequence 391 AA;
 Query Match 88.0%; Score 1237; DB 22; Length 391;
 Best Local Similarity 69.6%; Pred. No. 4,1e-125;
 Matches 250; Conservative 0; Mismatches 1; Indels 108; Gaps 1;
 QY 1 MAPELLQVGLRFRFOLHLVFILPSSMLLPIHLLENFAAATPGHRCWYHMDNNTGS 60
 DB 27 maeellsqvgvlgrrfgmhlvflfplsmlllphllenfaaapghrcwvhmdntgcs 86
 QY 61 GNETGIISEDALLRISIPDLSNLRPEKCRRYHPQWOLLHNGTHTSHSEADTEECYDVG 120
 DB 87 gnetgllsedallrlrlspldslndlrpekcrtrvhpqwlhngtthshsdeadtepcvqgw 146
 QY 121 VYDOSFPESTIVTKMDLVCDYQSLKSVQFLLLTGMLVGIGTGHVSD----- 168
 DB 147 vydgysfptslvckwldvcdyqslksvqfllltgmlvvgigghvsdrfgrrllrwgl 206
 QY 169 ----- 168
 DB 207 lqalttcaafaptfpyrcvlrflagfsmllsnnslptewlrpnkskalvvlssga 266
 QY 169 -----RMIVESARWLIITNRKIDEGIKALR 192
 DB 267 lnlggllygllyvfrdwqlhvasvpffvffllslrwlvsaarwlltnkldeglkalr 326
 QY 193 KYARNGIKNAEETINIEVVRSTMOEELDAQTKTTCVCDLFNPSMKRRCILVFLRK 251
 DB 327 kwarnglknaeetlnievvrstmqeeldaqtkttvcdlfnpsmrkrilvflrkk 385
 RESULT 4
 AAM24188
 ID AAM24188 standard; Protein; 213 AA.
 XX
 AC AAM24188;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Human EST encoded protein SEQ ID NO: 1713.
 XX
 KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KW diagnostics; forensic test; gene mapping; genetic disorder;
 KW biodiversity; gene therapy; nutrition.
 XX
 OS Homo sapiens.
 XX
 PN WO200154477-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 25-JAN-2001; 2001WO-US02687.
 XX
 PR 25-JAN-2000; 2000US-0491404.
 PR 17-JUL-2000; 2000US-0617746.
 PR 03-AUG-2000; 2000US-0631451.
 PR 15-SEP-2000; 2000US-0663870.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 PI Cao Y, Drmanac RA, Zhang J, Werlman T;
 XX
 XX WPI: 2001-476164/51.
 DR N-PSDB; AAH98847.
 XX
 PT Isolated polypeptide for treatment of diseases, diagnostics, raising
 antibodies and research use -

XX PS Claim 20; Page 1141; 1275pp; English.

XX CC The present invention provides the protein and coding sequences of novel

CC proteins from a variety of organisms, including human, dog, cat, horse,

CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea

CC urchin and tomato. These were derived from expressed sequence tags (ESTs)

CC from the organism of interest. They can be used in diagnostics,

CC forensics, gene mapping, identification of mutations, to assess

CC biodiversity and for nutritional purposes. The present sequence is a

CC protein of the invention.

XX SO Sequence 213 AA:

Query Match 70.6%; Score 993; DB 22; Length 213;

Best Local Similarity 100.0%; Pred. No. 4.4e-99;

Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 MHLVFLPSLMLLPHILLENFAPGRCVHMLDNTGSGNETGISEDALTRIS 77

Db 1 mlhvfllpslmlllphlllenfaaiphrcv/hmlhntgsgnetglsedallrissi 60

QY 78 PLDSNLREPKCRFRVHPQQLHLNCTHSTSEADTEPCVDGVTYQSTPSTIVTKWDL 137

Db 61 pldsnlrepekcrfrvhpqqlhlhngtlhntseadtepcvdgvtvdyqsyfprstivtkwdl 120

QY 138 VCDYQSLKSVQFLLLTGMLVGGIIGHVSDRWLVESARWLITTNKLDKGLKLRVART 197

Db 121 vcdyqslksvqfllltgmlvggiighvsvdrwlvesarwllitnkldegiklrvart 180

QY 198 NGIRNAE 204

Db 181 ngirknae 187

RESULT 5

AAE04897

ID AAE04897 standard; Protein: 553 AA.

XX AC AAE04897;

XX DT 10-SEP-2001 (first entry)

XX DE Human transporter and ion channel-10 (TRICH-10) protein.

XX KW Human: transporter and ion channel-10; TRICH-10; cystic fibrosis; mood;

KW gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy;

KW hypertension; angina; neurological disorder; asthma; bipolar disorder;

KW dementia; depression; Alzheimer's disease; epilepsy; vaccine; arrhythmia;

KW Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke;

KW Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy;

KW demyelinating disease; mental disorder; Schizophrenia; polymyositis;

KW dermatomyositis; diabetes mellitus; cataract; myocarditis; Grave's disease;

KW rheumatoid arthritis; Sjogren's syndrome; immunological disorder; psoriasis;

KW sickle cell anaemia; Wilson's disease; systemic lupus erythematosus;

KW scleroderma; pulmonary artery stenosis; nontropic; Addison's disease;

KW malabsorption syndrome; hypercholesterolaemia; cancer.

XX OS Homo sapiens.

XX FH Key

XX FH Location/Qualifiers

FT Domain 204..222

FT Domain /label= Transmembrane_domain

FT Domain 470..493

FT Domain /label= Transmembrane_domain

FT Domain 500..519

FT Domain /label= Transmembrane_domain

XX PN WO200146258-A2.

XX PD 28-JUN-2001.

XX PF 22-DEC-2000; 2000WO-US35095.

XX PR 23-DEC-1999; 99US-0172000.

PR 14-JAN-2000; 2000US-0176083.

PR 21-JAN-2000; 2000US-0177332.

PR 28-JAN-2000; 2000US-0178372.

PR 02-FEB-2000; 2000US-0179758.

PR 10-FEB-2000; 2000US-0181625.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Baughn MR, Burford N, Au-Young J, Lu DAM, Yang J, Reddy R, Lal P;

PI Hillman JU, Azimzal Y, Yue H, Nguyen DB, Yao MG, Gandhi AR;

PI Tang YT, Khan FA;

XX DR WPI: 2001-418042/44.

XX DR N-PSDB; AAD09561.

PT Novel human transporter and ion channel proteins useful for treating

PT disorders

PT disorders

XX PS Claim 1; Page 122-123; 160pp; English.

XX CC The present sequence is transporter and ion channel-10 (TRICH-10)

CC protein. TRICH is used as vaccine. TRICH is useful for treating a disease

CC or condition associated with decreased expression of functional TRICH,

CC such as transport disorder including amyotrophic lateral sclerosis,

CC cystic fibrosis, Becker's muscular dystrophy, Charcot-Marie-Tooth

CC disease, Duchenne muscular dystrophy, angina and hypertension,

CC neurological disorders including Alzheimer's disease, amnesia, bipolar

CC disorder, dementia, depression, epilepsy, ischaemic cerebrovascular

CC disease, stroke, cerebral neoplasms, Pick's disease, Huntington's

CC disease and Parkinson's disease, demyelinating diseases, mental disorders

CC including mood, anxiety, Schizophrenia and seasonal affective disorder,

CC muscle disorder including cardiomyopathy, myocarditis, polymyositis,

CC dermatomyositis, arrhythmias and asthma and immunological disorders,

CC including AIDS, adult respiratory distress syndrome (ARDS), allergies,

CC anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's

CC syndrome, systemic lupus erythematosus and other diseases including

CC sickle cell anaemia, Wilson's disease, cataracts, infertility, pulmonary

CC artery stenosis, Grave's disease, Cushing's disease, Addison's disease,

CC glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers

CC psoriasis and viral, bacterial, fungal, helminthic and protozoal

CC infections. TRICH DNA is useful in gene therapy and in diagnostic

CC purposes.

XX SO Sequence 553 AA:

Query Match 56.9%; Score 799.5; DB 22; Length 553;

Best Local Similarity 44.7%; Pred. No. 1.7e-77;

Matches 160; Conservative 40; Mismatches 49; Indels 109; Gaps 2;

QY 1 VAFEEILSOVGGIGRQMLHLVFLPSLMLLPHILLENFAPGRCVHMLDNTGSG 60

Db 1 mafgdlignagdlwrfqltqvflsfavatyhmlentafipgrcvhllndtvs 60

QY 61 GNETGISEDALTRISIPDSNLREPKCRFRVHPQQLHLNCTHSTSEADTEPCVDGW 120

Db 61 dntgtalsgdalltrissipdsnlrepekcrfrvhpqqlhlhngtlfptsdamepcvdgw 120

QY 121 VYDQSTFPTIVTKWDLVCDYQSLKSVQFLLLTGMLVGGIIGHVSDR----- 169

Db 121 vydrsfstltvewdlvcdsgslsvakfvfmgmvggllgghlstdtgrtfrvltwcy 180

QY 170 ----- 169

Db 181 lqvavtcoaalapfllycsirfsglaeaslltntlmiaewatbrqamgtlmgcp 240

QY 170 -----WLVESARWLITTNKLDKGLKLR 192

Db 241 sgiafmltaglafairdwhlqlvsvpyfvlftslswllesarwllinnkpeegjkeir 300
 QY 193 KVARNGIKNAEETINIEVVRSTMOEELDAQ-TKTTVCDFRNPSMKRRICTIVFLR 249
 301 kaahrsgmknardltlletlksctmkleaaqkkpslcmhmpnckrislstr 358

RESULT 6
 AAB47275
 ID AAB47275 standard; Protein: 554 AA.

AC AAB47275;

DT 06-AUG-2001 (first entry)

DE hOAT4.

Human; organic anion transporter; hOAT; liver; kidney;
 membrane protein; transport; organic anion; splice variant.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 324 /note="PKC phosphorylation site"

PN MO200104283-A2.

PD 18-JAN-2001.

PF 12-JUL-2000; 2000MO-US18980.

PR 12-JUL-1999; 99US-0143771.

XX (META-) METABASIS THERAPEUTICS INC.

PI Sun W;

XX WPI: 2001-367057/38.

DR N-PSDB: AAC85823.

XX Nucleic acids encoding human organic anion transporter polypeptides,
 useful in gene therapy procedures -

PT Claim 2; Fig 5; 95pp; English.

XX The sequences given in AAB47271-76 represent human organic anion
 transporter (hOAT) polypeptides. hOAT polypeptides are preferentially
 expressed in the liver and kidneys of humans. OAT's are membrane
 proteins that facilitate the transport of organic anions across the
 cell membrane. The mechanism of transport is thought to be a secondary
 or tertiary active transport involving exchange of another organic anion.
 hOAT2A and hOAT2B are thought to be splice variants as they are
 identical except at the C-terminal end. hOAT proteins and the DNA
 encoding them, may be used in the prevention, treatment and diagnosis
 of diseases associated with inappropriate hOAT expression.

CC Sequence 554 AA:

XX SQ

Query Match 54.6%; Score 768; DB 22; Length 554;
 Best Local Similarity 43.7%; Pred. No. 4,4e-74;

Matches 157; Conservative 40; Mismatches 52; Indels 110; Gaps 3;

QY 1 MAREELLGQVGGIRFOMLHVFILPSMLLPHILENFAAIPGRCVWMLDNNTGS 60

Db 1 matgdlighgdltwriglqgtvlstlftavatyhfmntafatipghrcwvhlldnlys 60

QY 61 GNETGILSEDALLRISIPLDNSNRPEKCRFVHPQQLHLNNTISTSPADPEPCVDGW 120

Db 61 dndtgalsgdallrltislpldsnmrpekcrtfvpwqlhnlnglftutsdaamepcvdgw 120

QY 121 VYDSYFPPTI-VTKMWLVCDYQSLKSVQFLTLTGLVGIIGHVSDR----- 169

Db 121 yldrlsfstldlkwldvcdsgsltsvakfvmagmmllgdlghlsdrgrsfvrlwrc 180

QY 170 ----- 169

Db 181 ylgvalvgcaalaptfllycsvrfllsgiaamsflntntlmllaewathrfgamgtlqmc 240

QY 170 ----- 191

Db 241 psgiafmltaglafairdwhlqlvsvpyfvlftslswllesarwllinnkpeegjkeir 300

QY 192 KVARNGIKNAEETINIEVVRSTMOEELDAQ-TKTTVCDFRNPSMKRRICTIVFLR 249

Db 301 rkaahrsgmknardltlletlksctmkleaaqkkpslcmhmpnckrislstr 359

RESULT 7

AAM40059
 ID AAM40059 standard; Protein: 341 AA.

AC AAM40059;

DT 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 3204.

DE Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 peripheral nervous system neuropathy; central nervous system; CNS;

XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

XX leukaemia.

OS Homo sapiens.

XX WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000MO-US34263.

PR 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 Wang Z, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 Zhao QA, Zhou P, Goodrich R, Drmanac RT;

DR WPI: 2001-442253/47.

XX N-PSDB: AAI59215.

XX Novel nucleic acids and polypeptides, useful for treating disorders
 such as central nervous system injuries -

XX Example 5; SEQ ID NO 3204; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
 the encoded polypeptides (AAM38642-AAM4213) with nootropic,

XX immunosuppressant and cytostatic activity. The polynucleotides are useful
 in gene therapy. A composition containing a polypeptide or polynucleotide
 of the invention may be used to treat diseases of the peripheral nervous
 system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 341 AA;

Query Match 54.4%; Score 765; DB 22; Length 341;
 Best Local Similarity 45.5%; Pred. No. 4.5e-74;
 Matches 152; Conservative 33; Mismatches 41; Indels 108; Gaps 1;

QY 1 MAFELLISQVGLGRFQMLHVFILPSLMLLPHILENFAAIPGRCWHMDNNTGS 60
 Db 1 mafgdlighagdlwrfqllqvflsfavatyhlhmlentafipgrcwvhlldndtvs 60
 QY 61 GNETGILSEDLALRISIPIDSNLRPEKCRFVHPQOMLHLNGTIHSTSEADTEPCVDGW 120
 Db 61 dndtgalsgdallrlstipldsmmrpekrrfvpqwlhlhngtlfptsdadmepcvdgw 120
 QY 121 VYDOSYFSPSTIYTKMDLVCDYQSLKSVQFLLTGMVGGIIGGHSR 169
 Db 121 vydrisfsaisvewdlvcdsgsltsvakfvfmgamvvgllgshdsrffrrfvlwcy 180
 QY 170 ----- 169
 Db 181 lqvavgtcaalapflfilyclrlfsgiaamslntnmlleawathrfqangitlmgcp 240
 QY 170 ----- 169
 Db 241 sgiafmltaglafalrdwhllqlvsvpyfvflfleswllsearvllnnkpeegikrlr 192
 QY 193 KVARNGIKNAEETLINIEVVRSTMOEELDAQTK 226
 Db 301 kahtsgmknardtcltleikstmkleaaqkk 334

RESULT 8

AAE19406
 ID AAE19406 standard; Protein; 339 AA.

AC AAE19406;

DT 06-MAR-2001 (first entry)

DE Amino acid sequence of a human secreted protein.

KW Secreted protein; platelet disorder; stem cell disorder; osteoporosis;
 KW osteoarthritis; burn; incision; ulcer; periodontal disease; neuropathy;
 KW nervous system disease; bone growth; cosmetic plastic surgery;
 KW gut protection; gut regeneration; fibrosis; cancer;
 KW bone marrow transplantation.

OS Homo sapiens.

PN WO200061755-A2.

PD 19-OCT-2000.

PF 10-APR-2000; 2000MO-US09555.

PR 09-APR-1999; 99US-0128574.

PR 20-AUG-1999; 99US-0150054.

PA (CHIR) CHIRON CORP.

PI Garcia PD;

DR WPI; 2000-665133/64.

DR N-PSDB; AAC61892.

XX Novel secreted human proteins useful for stimulating blood cell
 PT generation in patients receiving cancer chemotherapy, treating bone
 PT marrow transplantation patients and for healing fractured bones
 XX Claim 14; Page 72; 74pp; English.

AAE19393-B19407 represent secreted human proteins. The secreted proteins
 CC are useful in assays to determine their biological activities. The
 CC proteins can also be used as biomarkers to identify tissues or cell
 CC types which express the proteins. The polynucleotide molecules or cell
 CC used as biomarkers for tissues or chromosomes and to elicit immune
 CC responses. The proteins and antibodies are useful in diagnosis and
 CC treatment of diseases associated with altered expression of these
 CC proteins. The proteins are also useful for prevention or treatment of
 CC platelet disorders, stem cell disorders, osteoporosis or osteoarthritis,
 CC burns, incisions, ulcers, periodontal diseases, central and peripheral
 CC nervous system diseases and neuropathies, for healing fractured bones
 CC and to induce cartilage and/or bone growth in cosmetic plastic surgery.
 CC The proteins are also useful for gut protection or regeneration, for the
 CC treatment of lung or liver fibrosis, for stimulating blood cell
 CC generation in patients receiving cancer chemotherapy and for treatment
 CC of bone marrow transplantation patients.

XX Sequence 339 AA;

Query Match 54.0%; Score 759; DB 21; Length 339;
 Best Local Similarity 44.9%; Pred. No. 2e-73;
 Matches 151; Conservative 33; Mismatches 44; Indels 108; Gaps 1;

QY 1 MAFELLISQVGLGRFQMLHVFILPSLMLLPHILENFAAIPGRCWHMDNNTGS 60
 Db 1 mafgdlighagdlwrfqllqvflsfavatyhlhmlentafipgrcwvhlldndtvs 60
 QY 61 GNETGILSEDLALRISIPIDSNLRPEKCRFVHPQOMLHLNGTIHSTSEADTEPCVDGW 120
 Db 61 dndtgalsgdallrlstipldsmmrpekrrfvpqwlhlhngtlfptsdadmepcvdgw 120
 QY 121 VYDOSYFSPSTIYTKMDLVCDYQSLKSVQFLLTGMVGGIIGGHSR 169
 Db 121 vydrisfsaisvewdlvcdsgsltsvakfvfmgamvvgllgshdsrffrrfvlwcy 180
 QY 170 ----- 169
 Db 181 lqvavgtcaalapflfilyclrlfsgiaamslntnmlleawathrfqangitlmgcp 240
 QY 170 ----- 169
 Db 241 sgiafmltaglafalrdwhllqlvsvpyfvflfleswllsearvllnnkpeegikrlr 192
 QY 193 KVARNGIKNAEETLINIEVVRSTMOEELDAQTKT 228
 Db 301 kahtsgmknardtcltleikstmkleaaqkk 336

RESULT 9

AAE10332
 ID AAE10332 standard; Protein; 553 AA.

AC AAE10332;

DT 10-DEC-2001 (first entry)

DE Human transporter and ion channel-9 (TRICH-9) protein.

KW Human; transporter and ion channel; TRICH-9; therapy; akinesia; cardiac;
 KW neurological disorder; immune disorder; allergy; nocturnal; dementia;
 KW AIDS; Acquired Immune Deficiency Syndrome; amyotrophic lateral sclerosis;
 KW cystic fibrosis; Alzheimer's disease; depression; epilepsy; scleroderma;
 KW cardiomyopathy; muscular dystrophy; Crohn's disease; multiple sclerosis;
 KW rheumatoid arthritis; cancer; anaemia; asthma; trauma; haemodialysis;
 KW gastritis; inflammation.

Dd 181 lqmwavgtaafapafpyclfflflafavayvmmntgtllnewtaatarplvmntlnslg 240
Oy 170 -----WVESARWLITNNKLDGLKALR 192
Dd 241 fsfghgltaavaygyrdwllqlvsvvpfficflyswbleasarvlltgcgrldwgldew 300
Oy 193 KVARFNGCIRKAEEFTLNIEVVRSTNOEELDAAKOTRYCDLFRNPSMKRRICI 244
Dd 301 rvaalngkxgvgdlipervlleamreelsmgppaslglltmptlfrfrci 352

RESULT 10
AA041845
ID AA041845 standard; Protein; 369 AA.
XX
AC AAM41845;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6776.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoclastic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

XX
OS Homo sapiens.
XX
PN WO200153312-A1.
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000MO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.

(HYSE-) HYSEQ INC.
XX
PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang Y, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

WPI: 2001-442253/47.
N-PDSB: AA161001.
DR Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PT
XX Example 2; SEQ ID NO 6776; 10078pp; English.
PS
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AA42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotide
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemoclastic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 369 AA;

Query Match
 Best Local Similarity 42.3%; Score 594.5; DB 22; Length 369;
 Matches 123; Conservative 30; Mismatches 32; Indels 109; Gaps 2;

QY 65 GILSEDLALRISIPDSNLRPEKCRREVHOMQLHNGTISTSEADTEPCVDGWYDQ 124
 Db 9 galsqadallrlstipdsnmpekcrtivdpqwlhngtftptsdadmpcvdgvwydr 68
 QY 125 STPTSTVTKWDVCDYQSKSVYQPLLTGMVGGIIGCHVSDR----- 169
 Db 69 lsfstltvewdvcvdsqslsvaktfvfmagmvgllghlsdtrgrfvrlwcylyva 128
 QY 170 ----- 169
 Db 129 lvtcaalaprfllycsrlflsgiaamslltntlllaewathrfqamgiltmgpsgia 188
 QY 170 ----- 169
 Db 189 fmlaglaflrtdwhllglvsvpyvflfllcswlllesarwlhnnpeegiketrkaah 196
 QY 197 TNGIKNAEETINIEVNSTQMOEELDAQTKTT-VCDFLRNPMKRRCILVFLR 249
 Db 249 rsgmknardltlrlkstkmlkelaqaqkkpflgerlmpnckrtislrlpfk 302

RESULT 11

ID AAE06612 standard; Protein: 550 AA.
 AC AAE06612;

25-SEP-2001 (first entry)

Human protein having hydrophobic domain, H030882.

KW Human; hydrophobic domain; gene therapy; nutritional supplement;
 KW multiple sclerosis; immunomodulatory; autoimmune disorder; antimicrobial;
 KW haematopoiesis; tissue growth activity; Parkinson's disease; diabetes;
 KW Huntington's disease; Alzheimer's disease; chemotactic; cytostatic;
 KW haemostatic; thrombolytic; tumour growth inhibitor; anabolic;
 KW contractile; antinflammatory; antinflammatory.

OS Homo sapiens.

PN WO200149728-A2.

PD 12-JUL-2001.

PF 28-DEC-2000; 2000WO-JP09359.

PR 06-JAN-2000; 2000JP-0000585.

PR 06-JAN-2000; 2000JP-0000588.

PR 03-FEB-2000; 2000JP-0002239.

PR 03-MAR-2000; 2000JP-0026862.

PA (PROT-) PROTEGENE INC.

PI (SAGA) SAGAMI CHEM RES CENT.

DR Kato S, Kimura T;

DR WPI: 2001-418355/44.

DR N-PSDB; AADI2607.

PT Human proteins with hydrophobic domains and the nucleic acids encoding
 PT them, useful for preventing diagnosing and treating e.g. cancer.

PT Alzheimer's and inflammation -

PS Claim 1; Page 137-138; 563pp; English.

CC The present sequence is human protein with hydrophobic domain,
 CC H030882. The polynucleotide and polypeptide of the invention
 CC may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate polypeptide expression. The polynucleotides
 CC may be used to produce the polypeptide, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the protein. The
 CC polynucleotides and its complementary sequences may also be used as DNA
 CC probes in diagnostic assays and also used in gene therapy. The
 CC and in assays to identify modulators of polypeptide expression and
 CC activity. The polypeptides and nucleic acids may be used as nutritional
 CC supplements, to modulate cytokine and cell proliferation activity, to
 CC modulate immune stimulation or suppression (e.g. for the treatment of
 CC microbial infections and autoimmune disorders such as multiple sclerosis,
 CC rheumatoid arthritis and insulin-dependent diabetes), to modulate
 CC treatment of Parkinson's disease, Huntington's disease and Alzheimer's
 CC disease), to modulate activin and inhibin activity (e.g. for controlling
 CC fertility), to modulate chemotactic and chemokinetic activity, to
 CC modulate haemostatic and thrombolytic activity, to modulate receptor
 CC ligand activity, to modulate inflammation and to inhibit tumour growth.

Sequence 550 AA.

Query Match
 Best Local Similarity 42.1%; Score 592; DB 22; Length 550;
 Matches 143; Conservative 28; Mismatches 69; Indels 116; Gaps 7;

QY 1 MAFEDILSQVGLGFGFQWLH-LVFIPLSLMLIP-HILLENFALRGHCQWMLDNT 58
 Db 1 mafeklllegagvgylfqllytlilpdlm--lpsgmllenfesaalpgbhwctmldn-- 56
 QY 59 GSGNETGILSEDLALRISIPDSNLRPEKCRREVHOMQLHNGTISTSEADTEPCVD 118
 Db 57 gsaavstn-mrpkallstlslpppnpqghqgrtrfrpqwqlldpnatatswseadtepcvd 115
 QY 119 GWYDOSYFPTSTVTKWDVCDYQSKSVYQPLLTGMVGGIIGCHVSDR----- 169
 Db 116 gwyydsvfsvtlvawdvcvdsqslkplsgslfmsgllvsgffwllsyrgrkpmjsw 175
 QY 170 ----- 169
 Db 176 colqlavagstlflaprtfviyglirfaafmagflslsltmewttsrravltmvg 235
 QY 170 ----- 169
 Db 236 calsagaalglafalrdwtllqlaasvpffalslswllpsarawllykypddgalge 295
 QY 191 LKRVARTNGIKNAEETINIEVNSTQMOEELDAQTKTYVCDLFPNPMKRRCILV 246
 Db 296 lrvkarlngkhea-knltlelvmsvkeavasakprsvldlfcvplvtrscaml 350

RESULT 12

ID AAB69091 standard; Protein: 550 AA.
 AC AAB69091;

24-APR-2001 (first entry)

Human organic anion transporter OAT4 protein sequence SEQ ID NO:2.
 Human organic anion transporter; OAT4; nephrotoxic; kidney disease;
 abnormal foetal growth.

Homo sapiens.

PN WO200102562-A1.
 PD 11-JAN-2001.
 XX 15-JUN-2000; 2000WO-JP03878.
 XX 01-JUL-1999; 99JP-0187244.
 XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX Endou H, Sekine T, Cha SH;
 PI WPI: 2001-138139/14.
 DR N-PSDB: AAF32614.
 DR Placental organic anion transporter gene and encoded polypeptide OAT4,
 XX useful in studying causes of abnormality, and in developing drugs to
 PT prevent or treat various kidney diseases and abnormal fetal growth
 PS Claim 2; Page 26-28; 32pp; Japanese.
 XX The present sequence represents a human placental organic anion
 CC transporter, designated OAT4. OAT4 has nephroretrophic activity. The OAT4
 CC protein and encoded gene are useful in studying causes of abnormality,
 CC including the application of a variant nucleic acid as a probe to detect
 CC the presence of a gene encoding the transporter OAT4, or to identify or
 CC quantify such gene, and in developing drugs to prevent or treat various
 CC kidney diseases and abnormal foetal growth.
 CC
 XX Sequence 550 AA:
 SQ
 Query Match 42.1%; Score 592; DB 22; Length 550;
 Best Local Similarity 40.2%; Pred. No. 5e-55;
 Matches 143; Conservative 28; Mismatches 69; Indels 116; Gaps 7;
 1 MAFEELLSQVGLGRFQMLH-LVFILPSLMLLP-HILLENFAAIFGHCWVHMDNNT 58
 1 mafeellseqagvgylftqlvltfllpclm--lpsqmllefsaiphwcwlmldn-- 56
 59 GSGNETGISSEDLARISFPLDSNLRPEKCRFRVHPQMLHNGTHTSSEATEPCVD 118
 59 gsgnetgisssedlarisfpldsnlrpekcrfrvhpqmlhngthtsseatpcvd 115
 57 gsavstn-mtkpalltisipppngphqcrfrfqpqqlldpnatatsweadepcvd 169
 119 GWYDSTPESTIVTKWDIVCDYQSLKSVYQFLTLTGMVGLIGHSVSDR----- 169
 116 gwyydstpestivtkwdivcdyqslksvyqfltlgtmvgllvgsfingllsytrfgkpmjsw 175
 170 ----- 169
 176 cclqlavagstlftapftvlycgllrfvaafmagiflssltlmwewlttsrravcmtyvg 235
 170 ----- 190
 236 cefsgaaglaiglafrdwrclqlaasvffafsliswlpesarwlilkykpdqalge 295
 191 LKPKVARTGINKNAEETLNIEVVRSTMOEELDAQTKTTVODLRNPSMRKRICILV 246
 236 lkpkvartginknaeetlnievvvrstmoeeeldaqtkttvoddrrnpsmrkricilv 246
 296 lkpkvartginknaeetlnievvvrstmoeeeldaqtkttvoddrrnpsmrkricilv 350
 296 lkpkvartginknaeetlnievvvrstmoeeeldaqtkttvoddrrnpsmrkricilv 350
 RESULT 13
 ID AAE10336 standard; Protein; 566 AA.
 XX AAE10336;
 AC AAE10336;
 XX 10-DEC-2001 (first entry)
 XX Human transporter and ion channel-13 (TRICH-13) protein.
 XX Human transporter and ion channel-13 (TRICH-13) therapy; akinesia;
 XX Human; transporter and ion channel; TRICH-13; therapy; nootropic; dementia;
 KW neurological disorder; immune disorder; allergy; nootropic; dementia;

XX AIDS: Acquired Immune Deficiency Syndrome; amyotrophic lateral sclerosis;
 KW cystic fibrosis; Alzheimer's disease; depression; epilepsy; scleroderma;
 KW cardiomyopathy; muscular dystrophy; Crohn's disease; multiple sclerosis;
 KW rheumatoid arthritis; cancer; anaemia; asthma; trauma; haemodialysis;
 KW gastritis; cardiac; inflammation.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 XX Region 103..543
 XX Domain /label= Sugar_transporter
 XX /label= 148..165
 XX /label= Transmembrane_domain

WO200162923-A2.
 PD 30-AUG-2001.
 XX 23-FEB-2001; 2001WO-US05942.
 XX 25-FEB-2000; 2000US-0184866.
 XX 02-MAR-2000; 2000US-0187947.
 XX 09-MAR-2000; 2000US-0188333.
 XX 17-MAR-2000; 2000US-0190230.
 XX 24-MAR-2000; 2000US-0192077.
 XX 30-MAR-2000; 2000US-0193500.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Yue H, Tang YT, Lal P, Policky JL, Nguyen DB, Au-Young J, Yao MG;
 PI Khan FA, Walla NK, Gandhi AR, Tribouley CM, Patterson C;
 PI Thornton M, Greene BD, Hernandez R, Borowsky ML, Sanjanwala MS;
 DR N-PSDB: AAD17480.
 XX WPI: 2001-582050/65.
 XX Thirteen human transporters and ion channels (referred to as TRICH-1 to
 PT TRICH-13), useful in the diagnosis, treatment and prevention of
 PT disorders (e.g. akinesia), neurological, muscle or immunological
 PT disorders (e.g. allergies) -
 PS Claim 1; Page 119-120; 131pp; English.
 XX The present sequence is human transporters and ion channels
 CC (TRICH-13) protein. The TRICH DNA, protein and their agonist and
 CC antagonists are useful in the diagnosis, treatment and prevention of
 CC transporter disorders (akinesia, amyotrophic lateral sclerosis, cystic
 CC fibrosis), neurological (Alzheimer's disease, dementia, depression),
 CC epilepsy), muscle (e.g. cardiomyopathy, Becker's muscular dystrophy)
 CC immunological disorders (e.g. Crohn's disease, acquired immunodeficiency
 CC syndrome (AIDS), Crohn's disease, rheumatoid arthritis, scleroderma,
 CC multiple sclerosis), viral, bacterial, parasitic, protozoal and
 CC helminthic disorders, cancer, anaemia, asthma, trauma, haemodialysis,
 CC gastritis and inflammation.
 CC
 XX Sequence 566 AA:
 SQ
 Query Match 42.0%; Score 590.5; DB 22; Length 566;
 Best Local Similarity 39.0%; Pred. No. 7.6e-55;
 Matches 140; Conservative 32; Mismatches 72; Indels 115; Gaps 6;
 1 MAFEELLSQVGLGRFQMLH-LVFILPSLMLLP-HILLENFAAIFGHCWVHMDNNT 58
 1 mafeellseqagvgylftqlvltfllpclm--lpsqmllefsaiphwcwlmldn-- 56
 59 GSGNETGISSEDLARISFPLDSNLRPEKCRFRVHPQMLHNGTHTSSEATEPCVD 118
 59 gsgnetgisssedlarisfpldsnlrpekcrfrvhpqmlhngthtsseatpcvd 115
 57 gsavstn-mtkpalltisipppngphqcrfrfqpqqlldpnatatsweadepcvd 169
 119 GWYDSTPESTIVTKWDIVCDYQSLKSVYQFLTLTGMVGLIGHSVSDR----- 169
 116 gwyydstpestivtkwdivcdyqslksvyqfltlgtmvgllvgsfingllsytrfgkpmjsw 175

QY 170 ----- 169
 Db 176 cclqlavagstlftaptfviyqglrfvaafmaglflsltlmwewttsrravmtlvvg 235
 QY 170 -----WVESARMLITNKLDDEGIKA 190
 Db 236 calsaggaalaglafalrdwrtqlaasvptfaistlswlpsarvlllkpkpdgaige 295
 QY 191 LRKARINGKINAEETLNEVVRSTMOEELDAQTCTTVCDFRNPSMRKRICLIVFLR 249
 Db 296 lrkvaringhkeetecvylkvlmsvaveasakeprsvldlfcvplvtrscamlvkv 354

RESULT 14

AAE06571
 ID AAE06571 standard; Protein; 578 AA.

AC AAE06571;

XX 25-SEP-2001 (first entry)

DE Human protein having hydrophobic domain, HP03613.

KW Human: hydrophobic domain; gene therapy; nutritional supplement;
 KW cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;
 KW multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;
 KW haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;
 KW Huntington's disease; Alzheimer's disease; chemotactic; tumour growth inhibitor; anabolic;
 KW contractile; thrombolytic; tumour growth inhibitor; anabolic;
 KW XX
 OS Homo sapiens.

XX MO200149728-A2.

PD 12-JUL-2001.

XX 28-DEC-2000; 2000WO-JP09359.

PR 06-JAN-2000; 2000JP-0000585.

PR 06-JAN-2000; 2000JP-0000588.

PR 11-JAN-2000; 2000JP-0002299.

PR 03-FEB-2000; 2000JP-0026862.

PR 03-MAR-2000; 2000JP-0058367.

XX (PROT-) PROTEGENE INC.

PA (SAGA) SAGAMI CHEM RES CENT.

XX Kato S, Kimura T;

XX WPI; 2001-418355/44.

DR N-PSDB; AAD12566.

XX Human proteins with hydrophobic domains and the nucleic acids encoding

PT them, useful for preventing diagnosing and treating e.g. cancer,

PT Alzheimer's and inflammation -

XX Claim 1; Page 59-60; 563pp; English.

XX The present sequence is human protein with hydrophobic domain, HP03613.
 CC The polynucleotide and polypeptide of the invention may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate polypeptide expression. The polynucleotides may be used
 CC to produce the polypeptide, by inserting the nucleic acids into a host
 CC cell and culturing the cell to express the protein. The polynucleotides
 CC and its complementary sequences may also be used as DNA probes in
 CC diagnostic assays and also used in gene therapy. The polypeptides may
 CC also be used as antigens in the production of antibodies and in assays
 CC to identify modulators of polypeptide expression and activity. The
 CC polypeptides and nucleic acids may be used as nutritional supplements,
 CC to modulate cytokine and cell proliferation activity, to modulate
 CC immune stimulation or suppression (e.g. for the treatment of microbial

CC infections and autoimmune disorders such as multiple sclerosis,
 CC rheumatoid arthritis and insulin-dependent diabetes), to modulate
 CC haematopoiesis, to modulate tissue growth activity (e.g. for the
 CC treatment of Parkinson's disease, Huntington's disease and Alzheimer's
 CC disease), to modulate activin and inhibin activity (e.g. for controlling
 CC fertility), to modulate chemotactic and chemokinetic activity, to
 CC modulate haemostatic and thrombolytic activity, to modulate receptor
 CC ligand activity, to modulate inflammatory activity and to inhibit tumour growth.
 CC XX
 SQ Sequence 578 AA;

Query Match 41.9%; Score 588.5; DB 22; Length 578;
 Best Local Similarity 35.3%; Pred. NO. 1.3e-54;
 Matches 133; Conservative 30; Mismatches 81; Indels 133; Gaps 1;

QY 1 MAFBELLSQVGLGRFQMLHLVFLPSIMLIPHIILENFAAIPGRVWHLMDNNTGS 60
 Db 1 mafbellldlvglgrfvgqlqmalmslmwlcqsmleatsaavshrcwaplidsnlaq 60
 QY 61 GNETGILSEDAALRISIPLDNSNLREKCRFVHPQWOLHNLNGTISTSEADTEPCVDGM 120
 Db 61 aslgsispeallaisippngqphqcrtrfipqwglldpnatatsweadepcvagw 120
 QY 121 VPDOSYPSPTIYKKNLDVCDYOSLKSVOFLTLTGMVGGIIGHSR- 169
 Db 121 vpdysifstlvaakwlvcdshalkpmagsilylagllygaacspasdrfgrllylwsy 180
 QY 170 ----- 169
 Db 181 lqmayvmtaafapafpyclfrflafavagvmmntgltrsltwfhagylhgsraep 240
 QY 170 ----- 169
 Db 241 lglawmewtaaraaplvmctlnslgfsfghltaavaygyrdwcllqlvsvpfllcflly 300
 QY 170 --WVESARMLITNKLDDEGLKALRKVARTNGKINAEETLNEVVRSTMOEELDAQTKT 227
 Db 301 swlaesarvlltgrldwsgelwraalngvgaqgdltlpevllsammeeismsgppa 360
 QY 228 TVCDLFRNPSMRKRICI 244
 Db 361 slgtllmpglrftcl 377

RESULT 15

AAU32379

XX AAU32379 standard; Protein; 645 AA.

XX AAU32379;

XX 18-DEC-2001 (first entry)

DE Novel human secreted protein #2870.

XX Human: vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

XX WO200179449-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US08656.

XX 18-APR-2000; 2000US-0552929.

XX 26-JAN-2001; 2001US-0770160.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT;

XX PI

XX
DR WPI: 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy -

XX
PS Claim 20: Page 597; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.

XX
SQ Sequence 645 AA:

Query Match 33.6%; Score 472; DB 22; Length 645;
Best Local Similarity 28.8%; Pred. No. 6.3e-42;
Matches 126; Conservative 31; Mismatches 63; Indels 218; Gaps 6;

QY 25 LPSLMLLP-HILENFAAIPGHRGWVHMLDNNNGSGNETGILSEDAALRISIPDSNL 83
DB 1 lpcdm--lpsqmllefaaiaipghrcwchmdn--gsavstn-mtpkalltisipgpnuq 55
QY 84 RPEKCRFVHPQWQLLNHNGTHTSTSEADTEPCYDGVWYDQSYFPSTIYTKWMDLYCDYQS 143
DB 56 gphgcrtrfqpqwlldpnatatsweadtepcvdgwydrsvftslvakwdlycsssqg 115
QY 144 LKSVVQFLLTGMLVGGIIGHVSDR----- 169
DB 116 lkpisqslfmsgllvgsflwgllysfgrfkpmalswccqlavagsttfapctfvlcyqlr 175
QY 170 ----- 169
DB 176 fvaafgmagiflsjltlmewettsterravtmvvgcafsaggaalglafalrdwrtlgl 235
QY 170 -----WLVESARMLIITNKLDDEGLKALKRYARTNGIKNAE----- 204
DB 236 aasvpfaaislswlipesarwllkqkpdqalqelrkvaringhkeaknlltletppppp 295
QY 205 ----- 204
DB 296 plpipsptapjstpcitftaltspappipisltpgpplqnisttfttvlignstll 355
QY 205 -----ETLNIEVVRSTMOBELDAOTKTITVCD 231
DB 356 insthvtstlthlhthvarthimgvldvgelecaylklmsvkeevaseakeprsvld 415
QY 232 LFRNPSMRKRICILVFLR 249
DB 416 lfcpvplwrtscamlvkvk 433

Search completed: July 1, 2002, 11:39:06
Job time: 372 sec


```

Db 61 V-----LPLCPNGKPEKCLRFVHP-----NASLPRDQ 89
Qy 111 ADREPCVDGVNDQSYFPPSTIYTKMDLVCDSIKSVYQGLLTLGMLVGGIIGGVSDR- 169
Db 90 GATEPCLDGMWYNT--RDITVEMDLVCGSNKLEKMAQSVFAGLIVGSPVIGELSDRF 147
Qy 170 -----169
Db 148 GRKPLTWGTYLLAAGSSAAPSPLVYMIFFLCGCSIGISLSTIILNVEWPTSTR 207
Qy 170 -----WVESARMLITN 182
Db 208 AISTTIGCYITGFIPLGLAVAPQWRMLQLSVSAFFISLSMWVPESTRMLVLSG 267
Qy 183 KLDGLKALRKVARTNGIKNAEETINIEVYSTMOEELDAQTKTVCLFRNPSMRK 240
Db 268 KSRALKTLQRAVATNGKKEBEKLTVEELKFNLDKDTISAKVYGLSDLFVSLIR 325

```

RESULT 2

```

US-08-501-572-1
; Sequence 1, Application US/08501572
; Patent No. 6063623
; GENERAL INFORMATION:
; APPLICANT: Koepsell, Hermann
; APPLICANT: Grundeman, Dirk
; APPLICANT: Gorboulev, Valentin
; TITLE OF INVENTION: Transport protein which effects the
; TITLE OF INVENTION: Transport of cationic xenobiotics and/or pharmaceuticals,
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/501,572
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Toohey, Kimberlin M
; REGISTRATION NUMBER: 35,391
; REFERENCE/DOCKET NUMBER: 02481.1453-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 556 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-501-572-1

```

Query Match 12.1%; Score 169.5; DB 3; Length 556;
 Best Local Similarity 20.5%; Pred. No. 7.8e-11;
 Matches 77; Conservative 47; Mismatches 95; Indels 157; Gaps 15;

```

Qy 4 EELISQVGLGRFQ---MLHLVFLIPSLMLIPHLLENFAAIPGRHC-----49
Db 5 DDVLEQVGEFQKQAFLLCLISASLAPYGVIGFLGFT---FGHCQNGVAELSOR 61

```

```

Qy 50 -WYHMLDN-----TGSNETGILSEDALLRISIPDSNLRPEKCRFRVHPQMLHLN 102
Db 62 CGMSQAEELNTYVGLDPSSEAFSLSQ--CMRYEV--DMNQSTLDC--VDP-----LS 108
Qy 103 GIHTSTSEADTEPCVDGVNDQSYFPPSTIYTKMDLVCDSIKSVYQGLLTLGMLVGGIIT 162
Db 109 SLVANSRLPLPGCEHGWYDTP--GSSITYERFLVCGDMKRYDLFQSCVNLGFFLSLV 166
Qy 163 GGHVSDR-----169
Db 167 VGTIADPFGKLLCLLVTLTVSVSVGLTAVAPDYTNLFRLLQGMVSKGWSGYTLIT 226
Qy 170 -----WLY-ES 174
Db 227 EFVSGYRRTATLYQMAFTVGLVGLAGVAYALPDRMWQLAVSLPTFLLYWFEVPS 286
Qy 175 ARPLITTNKLDGLKALRKVARTNG--IKNAEETINIEVYSTMOEELDAQTKT-TVCDL 232
Db 287 PFWLSQKRTTRAVRTMEDIAQNGKVPAD-----LKMCLDE--DASEKRSFADL 338
Qy 233 FRNPSMRKRICILVFL 248
Db 339 FRPMLRKHTVLMYL 354

```

RESULT 3

```

US-09-040-444-1
; Sequence 1, Application US/09040444
; Patent No. 6063766
; GENERAL INFORMATION:
; APPLICANT: Koepsell, Hermann
; APPLICANT: Grundeman, Dirk
; APPLICANT: Gorboulev, Valentin
; TITLE OF INVENTION: Transport protein which effects the
; TITLE OF INVENTION: Transport of cationic xenobiotics and/or pharmaceuticals,
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,444
; FILING DATE: March 18, 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Connor, Steven P
; REGISTRATION NUMBER: 41,225
; REFERENCE/DOCKET NUMBER: 2481.1453-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 556 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-040-444-1

```

Query Match 12.1%; Score 169.5; DB 3; Length 556;
 Best Local Similarity 20.5%; Pred. No. 7.8e-11;
 Matches 77; Conservative 47; Mismatches 95; Indels 157; Gaps 15;

RESULT 4
 US-08-501-572-3
 Sequence 3, Application US/08501572
 Patent No. 6063623
 GENERAL INFORMATION:
 APPLICANT: Koepsell, Hermann
 APPLICANT: Grundeman, Dirk
 APPLICANT: Garbouliev, Valentin
 TITLE OF INVENTION: Transport Of Cationic Xenobiotics and/or Pharmaceuticals
 TITLE OF INVENTION: Transport Of Cationic Xenobiotics and/or Pharmaceuticals
 TITLE OF INVENTION: DNA Sequences Encoding It And Their Use.
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
 STREET: 1300 I Street, N.W., Suite 700
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentln Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/501,572
 FILING DATE:
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Toohey, Kimberlin M
 REGISTRATION NUMBER: 35,391
 REFERENCE/DOCKET NUMBER: 02481.1453-00000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)408-4400
 TELEFAX: (202)408-4400
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 555 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-501-572-3

[illegible]

RESULT 5
 US-09-040-444-3
 ; Sequence 3, Application US/09040444
 ; Patent No. 6063766
 ;
 ; GENERAL INFORMATION:
 ; APPLICANT: Koepsell, Hermann
 ; APPLICANT: Grundeman, Dirk
 ; APPLICANT: Gorboulev, Valentin
 ; TITLE OF INVENTION: Transport of cationic xenobiotics and/or Pharmaceuticals
 ; TITLE OF INVENTION: Transport Of Cationic Xenobiotics and/or Pharmaceuticals
 ; TITLE OF INVENTION: DNA Sequences Encoding It And Their Use.
 ; TITLE OF INVENTION: 6
 ; NUMBER OF SEQUENCE ADDRESSES:
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.
 ; STREET: 1300 I Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ;
 ; CURRENT APPLICATION NUMBER: US/09/040,444
 ; APPLICATION NUMBER: 41,225
 ; FILING DATE: March 18, 1998
 ;
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: O'Connor, Steven P
 ; REGISTRATION NUMBER: 41,225
 ; REFERENCE/DOCKET NUMBER: 2481.1453-01
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)408-4400
 ; TELEFAX: (202)408-4400
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 555 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide

Tue Jul 2 10:13:39 2002

us-09-674-235-1.ra1

US-09-040-444-3

Query Match
Best Local Similarity 10.8%; Score 152.5; DB 3; Length 555;
Matches 71; Conservative 40; Mismatches 109; Indels 139; Gaps 12;

11 GGLGRFQMLHVFILPSMLLIPHLLE-NPAAIPGHRG---VWMLDNNNGSGNGEIGI 66
12 GGEFFRQKMFLLALSAFAPIYGIYLFETPDHCRSPGVAELSLRG-----64
67 LSEDAALRISIP---LDSNLRPEKRRFVHPQWLHLNGTHTSHSEADTF-----PC 116
65 WSPAEEINLYTVPGPGAGESSPQCRRY-EVDMNOSTFD-CVDPLASLDITRSLPLGPC 122
117 VDMGVYDOSYFSTYTKMDLVCDYQSLKSVQFLLTGMLVGITIGHVSDR-----169
123 RDGMVYETP--GSSIVTEFNLVCANSMWMLDFQSSVANGFFIGSMISGYIADFRGRLCL 180
170 -----WLV-----172
181 LFTVLMAAGVLAISPTYTMILFRLQGLVSKAGMLIGYLLITEFYGRRTVTCIE 240
173 -----ESARLITTKLDEGL 188
241 YQAVYVGLVAVAYALPHRMLOFVALPNEFFLYWCIPSPRMLLSQNKNEAM 300
189 KALKRVARTGINKNAEETLNIEVYSTMOEELDAQTKTYCDLFRNSMKRRICILVE 247
301 RIHKIAKKNG-----KSLPASILQRLLEE-TGKKLNPFLDLVTPQIRKHTMILMY 353

RESULT 6
US-08-501-572-2
Sequence 2, Application US/08501572
Patent No. 6063623
GENERAL INFORMATION:
APPLICANT: Koepsell, Hermann
APPLICANT: Grundeman, Dirk
TITLE OF INVENTION: Transport of cationic xenobiotics and/or pharmaceuticals,
TITLE OF INVENTION: Transport of cationic xenobiotics and/or pharmaceuticals,
TITLE OF INVENTION: Transport of cationic xenobiotics and/or pharmaceuticals,
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESS: Finegan, Henderson, Parabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/501.572
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Toohy, Kimberlin M
REGISTRATION NUMBER: 35,391
REFERENCE/DOCKET NUMBER: 02481.1453-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)408-4400
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 553 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-501-572-2

Query Match
Best Local Similarity 10.4%; Score 146; DB 3; Length 553;
Matches 75; Conservative 34; Mismatches 117; Indels 142; Gaps 11;

4 EELLQVSGGLGFQMLHVFILPSMLLIPHLLENPAAIPGHRG---VWMLDNNNGS 60
5 DDLEQVSGSGWQKQAFLLICLSAFAPIYGIYLFETPDHCRSPGVAELSLRG-----63
61 GNEGTIISDALRISIPLDSNLRPE-----KRRFVHPQWLHLN-----GTHTST 108
64 -----WSPAEEINLYTVPGPGAGESSPQCRRY-EVDMNOSTFD-CVDPLASLDITRSLPLGPC 122
109 SEADTEPCVDGWYDOSYFSTYTKMDLVCDYQSLKSVQFLLTGMLVGITIGHVSDR-----169
114 SHLPLGRCQDQWYDIP--GSSIVTEFNLVCANSMWMLDFQSSVANGFFIGSMISGYIADFRGRLCL 180
169 R-----WLV-----172
172 RGRKCLLGLTVLVAASVGLMAFSPNMSMLFRLQGLVSKAGMLIGYLLITEFYGRRTVTCIE 240
170 -----ESARLITTKLDEGL 188
232 SRRVIAIYQMAFTVGLVALGLAYALPHRMLOFVALPNEFFLYWCIPSPRMLLSQNKNEAM 300
181 TNKLEGIKALKRVARTGINKNAEETLNIEVYSTMOEELDAQTKTYCDLFRNSMKRRICILVE 247
292 QKRTTEAIKIMDHIQKNG-KLPPADLKMLSLEEDYTEKL-----SPSFADLFTPTRLRK 345
241 RICILVEL 248
346 RTFILMYL 353

RESULT 7
US-09-040-444-2

Sequence 2, Application US/09040444
Patent No. 6063766
GENERAL INFORMATION:
APPLICANT: Koepsell, Hermann
APPLICANT: Grundeman, Dirk
TITLE OF INVENTION: Transport of cationic xenobiotics and/or pharmaceuticals,
TITLE OF INVENTION: Transport of cationic xenobiotics and/or pharmaceuticals,
TITLE OF INVENTION: Transport of cationic xenobiotics and/or pharmaceuticals,
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESS: Finegan, Henderson, Parabow, Garrett & Dunner, L.L.P.
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040.444
FILING DATE: March 18, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: O'Connor, Steven P
REGISTRATION NUMBER: 41,225
REFERENCE/DOCKET NUMBER: 2481.1453-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)408-4400
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO: 2:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-040-444-2

```

Query Match	10.4%;	Score 146;	DB 3;	Length 553;
Best Local Similarity	20.4%;	Pred. No. 3.9e-08;		
Matches	75;	Conservative 34;	Mismatches 117;	Indels 142;
				Gaps 11

```

QY      4 EELLSSVGGIGRRPQMLHLVFLPISMLLPHILLEPMAALPEGHRG--WVHMLDNNTGS 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      5 DDLTEVGGESGWFQKQAFILLCLLSAFAFICVIGVFLGTFPHDHCSQPEVIELSORCG-- 63
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      61 GNFTGISSEALLRISIPLDNSLRPE-----KCRFFHPQOMLLHN-----GTHTST 108
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      64 -----WSPAEELNTYVP---GLGPPAGEAFLGCRHY-EVDWQMSALSCVDPPLASLATNR 113
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      109 SEADTEPCVDGWNVDOSYFPESTIVTKMDLVCDYQSLKSVYQFLLLGMLVGSIIGGVSD 168
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      114 SHLPFGCCQGWNYDPR--GSSIVTEEFNLVCAQMSWKLDLFQSCILNAGFFPQSLGVGYFAD 171
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      169 R----- 169
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      172 RFRGRICLLGTVLVNAVSGVLMAFSPNYMSMLLFRLLQGLVSKGNMAGTLLITEFVSG 231
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      170 -----WLV-ESARWLLI 180
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      232 SRRTVALIMYQMAFTVGLVALTGLALYALPHRMWQLAVSLPTFLFLYVWCVPESPRLWLS 291
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      181 TNKLDGKALARKVARTNGIKNADELTINIEVNSTOEBELDAQTKTYCDLFRNPBMRK 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      292 QKRRTVEAIKIKTDHIAQKNG--KLPPADKLKMLSEBDVTEKL-----SPSFADLFETPRLRK 345
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      241 RICILVFL 248
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      346 RTFTLMWYL 353
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 8
 US-08-964-127-2
 Sequence 2, Application US/08964127
 Patent No. 627555
 GENERAL INFORMATION:
 APPLICANT: Grandearl, Andrew David John
 TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE
 TITLE OF INVENTION: MOLECULES
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows 95
 SOFTWARE: FASTSEQ for Windows Version 2.0b
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/964,127
 FILING DATE: 06-NOV-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Crews, Ph.D., L. Lee
 REGISTRATION NUMBER: P-43,567
 REFERENCE/DOCKET NUMBER: 07334/038001

TELECOMMUNICATION INFORMATION
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-964-127-2

```

Query Match          7.43; Score 104; DB 4; Length 520;
Best Local Similarity 18.28; Pred. No. 0.0025;
Matches 64; Conservative 33; Mismatches 92; Indels 162; Gaps 14.

QY 39 NFAAALPGHRCVHMLMDNNTGSGNETGISLSEDAALLRISIPLDSSNLREPKCRREYHQMQL 98
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 16 HYGAPPNNSGM-EQPPNMSGVSVAASALAAASRASTSDPS-----CSGRAPP----- 65

QY 99 IHLNGTTHSTSEADTPEPCVDGWYD-QSIFPSTIYKMLVCD--YQSLKSYQFLL----- 152
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 66 -----DENHCLKMDWYNGLPVLTNAGQWMDVCDLGQVILEQLEFLILGFA 112
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 153 -----LQGMVLGSI----- 161
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Db 113 SGYLELGPADPRGRRGIVLLTLGLVPCGCVGAAGSSTGVAALEFLGLGFLLAGVDLGV 172
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 162 -----LGGH-----VSDRW----- 170
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Db 173 YLMRELECDPTQRLRALVAGELVGVGGHFLGLALVSKDMRFLOKMTAPCILLFLFYWG 232
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 171 ---LVESARMLITTKLDGLKALKRVYARPNGLKNAEETINIEVSVSTQMEELDAQOTK 227
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 233 PGLFLESARMLITVKQIEEASQSLRIIAERN-RPHQMG-----EEAQCADDLE 282
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 228 IYVCDL-----FNPSPMRKRICILVELRKIKSRKRRKNDQYTKV 265
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 283 NCPLEPATSTSEFASLLNTRN--IWKNLILGLTFNFAHAIRH---CYQPV 328

```

```

1  US-09-496-692-2
2  RESULT 9
3  Sequence 2, Application US/09496692
4  Patent No. 6313271
5  GENERAL INFORMATION:
6  APPLICANT: Grandearl, Andrew David John
7  TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE
8  TITLE OF INVENTION: MOLECULES
9  NUMBER OF SEQUENCES: 17
10 CORRESPONDENCE ADDRESSES:
11 ADDRESSEE: Fish & Richardson P.C.
12 STREET: 225 Franklin Street
13 CITY: Boston
14 STATE: MA
15 COUNTRY: USA
16 ZIP: 02110-2804
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: Diskette
19 COMPUTER: IBM Compatible
20 OPERATING SYSTEM: Windows 95
21 SOFTWARE: FASTSEQ for Windows Version 2.0b
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: US/09/496,692
24 FILING DATE:
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER: 08/964,127
27 FILING DATE: 06-NOV-1997
28 ATTORNEY/AGENT INFORMATION:
29 NAME: Crews, Ph.D., L. Lee
30 REGISTRATION NUMBER: P-43,567

```

LENGTH: 286 amino acids
TYPE: amino acid
TOPOLOGY: linear

Page 6

MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-964-127-4

Query Match 7.0%; Score 98.5; DB 4; Length 286;
Best Local Similarity 17.8%; Pred. No. 0.0043;

Matches 54; Conservative 29; Mismatches 73; Indels 147; Gaps 11;

39 NFAAIPGHCWVMDLNTGSGNETGILSEDALLRISIPIDSRLPEKCRFPVHPOMQL 98
16 HYGAFPPNAGSG-EQPPNAGSVASALAAASRVATSTDP-----CSGFAP----- 65
99 LHLNGTHTSHSEADTEPCVDGWYD-QSYFSTIYTKMDLYCD--YQSLKSYVQFL--- 152
66 -----DENHCLKMDYNGLPVLTNNAGMDLVCDLGMOVYILQIILFILGFA 112
153 -----LTGMLVSGI----- 161
113 SGYFLGYPADRFGRGIVLLTGLVPGCGVGAAGSSTGVMALRFLGLAGVLDGV 172
162 -----IGGH-----VSDRW----- 170
173 YLMRELCDPTQRLRVLAGELVGVGHFLFLGLALVSKDMRFLORMITAPCILEFLFYGW 232
171 ---LYESARWLITNKLDGELKALKRYA---RTNGIKNAEETLINIEVVRSTMOEELDAAQ 224
233 PGLFLESARWLIVKROIENASVLRILAEARNRHGOMLGEA-----QEAOLDLE 282
225 TKT 227
283 SST 285

RESULT 12

US-09-496-692-4

Sequence 4, Application US/09496692

Patent No. 6313271

GENERAL INFORMATION:

APPLICANT: Grandearl, Andrew David John

TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE

TITLE OF INVENTION: MOLECULES

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: FastSeq for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/496,692

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/964,127

FILING DATE: 06-NOV-1997

ATTORNEY/AGENT INFORMATION:

NAME: Crews, Ph.D., L. Lee

REGISTRATION NUMBER: P-43,567

REFERENCE/DOCKET NUMBER: 07334/038001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 286 amino acids

type: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-09-496-692-4

Query Match 7.0%; Score 98.5; DB 4; Length 286;
Best Local Similarity 17.8%; Pred. No. 0.0043;

Matches 54; Conservative 29; Mismatches 73; Indels 147; Gaps 11;

39 NFAAIPGHCWVMDLNTGSGNETGILSEDALLRISIPIDSRLPEKCRFPVHPOMQL 98
16 HYGAFPPNAGSG-EQPPNAGSVASALAAASRVATSTDP-----CSGFAP----- 65
99 LHLNGTHTSHSEADTEPCVDGWYD-QSYFSTIYTKMDLYCD--YQSLKSYVQFL--- 152
66 -----DENHCLKMDYNGLPVLTNNAGMDLVCDLGMOVYILQIILFILGFA 112
153 -----LTGMLVSGI----- 161
113 SGYFLGYPADRFGRGIVLLTGLVPGCGVGAAGSSTGVMALRFLGLAGVLDGV 172
162 -----IGGH-----VSDRW----- 170
173 YLMRELCDPTQRLRVLAGELVGVGHFLFLGLALVSKDMRFLORMITAPCILEFLFYGW 232
171 ---LYESARWLITNKLDGELKALKRYA---RTNGIKNAEETLINIEVVRSTMOEELDAAQ 224
233 PGLFLESARWLIVKROIENASVLRILAEARNRHGOMLGEA-----QEAOLDLE 282
225 TKT 227
283 SST 285

RESULT 13

US-08-989-386-7

Sequence 7, Application US/08989386

Patent No. 5989860

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Hillman, Jennifer L.

APPLICANT: Guejter, Karl J.

APPLICANT: Corley, Neil C.

APPLICANT: Shah, Puryi

TITLE OF INVENTION: HUMAN ISOMERASE HOMOLOGS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/989,386

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0443 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

QY	135	MDLYCDYOSKTSVYOFLITLTGMLYGIIGHVSDRLVYESRWLITLTKLDEGK	-ALRK	19
		-----	-----	
Db	184	WPLLGLGFLPAILIIOCALP	-----	228
		-----	-----	
QY	194	VARIINGIKNABELNIEVYRSTWGEELDAOTKTVDLFPNPSMRKICILVFLR		249
		-----	-----	
	229	LMGEEDVADODIÖEKDESMRSOKROY	-----	277
		-----	-----	

Tue Jul 2 10:13:39 2002

us-09-674-235-1.rail

Page 9

Search completed: July 1, 2002, 11:39:26
job time: 172 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 11:39:30 ; Search time 11.68 seconds
(without alignments)
888.428 Million cell updates/sec

Title: US-09-674-235-1

Perfect score: 1406
Sequence: 1 MAFEELLQVSGIGRFQMLH.....KKIKSRKHKNDCTKTKTKF 268

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	216.5	15.4	557	1 OCN2_MOUSE	O920E8 mus musculus
2	195.5	13.9	557	1 OCN2_RAT	O70594 rattus norv
3	193.5	13.8	557	1 OCN2_HUMAN	O76082 homo sapien
4	100	7.1	524	1 GTR2_HUMAN	P11168 homo sapien
5	95	6.8	523	1 GTR2_MOUSE	P14246 mus musculus
6	93.5	6.7	526	1 C11A_MOUSE	O99482 mus musculus
7	93	6.6	522	1 GTR2_RAT	P12336 rattus norv
8	87	6.2	581	1 FKX_MOUSE	O61576 mus musculus
9	86.5	6.2	496	1 GTR3_CHICK	P28568 gallus gall
10	84.5	6.0	636	1 GYRB_THEMA	P77993 thermocoga
11	83.5	5.9	494	1 GTR3_BOVIN	P58352 bos taurus
12	83.5	5.9	494	1 GTR3_SHEEP	P47843 ovine aries
13	83.5	5.9	520	1 C11A_CAPI	P79153 capra hircu
14	82.5	5.9	490	1 GTR1_CHICK	P46896 gallus gall
15	82.5	5.9	535	1 GHT3_SCHPO	O92339 schizosacch
16	82	5.8	492	1 GTR1_HUMAN	P11166 homo sapien
17	81.5	5.8	400	1 GTR3_RABIT	O93562 oryctolagus
18	81.5	5.8	608	1 YD56_YEAST	O92339 schizosacch
19	81	5.8	520	1 C11A_PIG	P10612 sus scrofa
20	81	5.8	561	1 YGGA_PIG	P53166 saccharomyc
21	80.5	5.7	260	1 CDH_MYCTU	O50676 mycobacteri
22	80.5	5.7	521	1 C11A_HUMAN	P05108 homo sapien
23	80	5.7	359	1 MKR2_DROME	P49071 drosophila
24	80	5.7	492	1 GTR1_MOUSE	P17809 mus musculus
25	80	5.7	492	1 GTR1_RABIT	P13805 oryctolagus
26	80	5.7	492	1 GTR1_RAT	P1167 rattus norv
27	80	5.7	520	1 C11A_MESAU	O99482 mus musculus
28	80	5.7	757	1 YOF2_CAEEL	O09314 caenorhabdi
29	79.5	5.7	520	1 C11A_HORSE	O46515 equus caball
30	79	5.6	198	1 SDC4_HUMAN	P31431 homo sapien
31	79	5.6	364	1 YW1_MYCTU	O10503 mycobacteri
32	79	5.6	520	1 C11A_BOVIN	P00189 bos taurus
33	79	5.6	629	1 PLST_HUMAN	Q14651 homo sapien

34	78.5	5.6	495	1 GTR3_CANEA	P47842 canis famli
35	78	5.5	1148	1 A8A2_MOUSE	P68200 mus musculus
36	77.5	5.5	1124	1 PHX_MOUSE	P33523 mougeotia s
37	77	5.5	270	1 VSP1_ARATH	O49195 arabidopsis
38	77	5.5	526	1 C11A_RAT	P14137 rattus norv
39	77	5.5	615	1 SPPA_HAERIN	P45243 haemophilus
40	77	5.5	1693	1 R1P2_MOUSE	P97433 mus musculus
41	76.5	5.4	201	1 YPT1_PHTIN	O01890 phytophthor
42	76.5	5.4	496	1 GTR3_HUMAN	P11169 homo sapien
43	76.5	5.4	1420	1 APX_XENLA	O01613 xenopus lae
44	76	5.4	315	1 VN34_ROTSA	O00721 stigmatia 11 r
45	76	5.4	376	1 Y433_MYCTU	P96279 mycobacteri

ALIGNMENTS

RESULT 1
OCN2_MOUSE STANDARD; PRT; 557 AA.
ID OCN2_MOUSE
AC O920E8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Organic cation/carnitine transporter 2 (Solute carrier family 22, member 5) (High-affinity sodium-dependent carnitine cotransporter).
GN SLC22A5 OR OCTN2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_Taxid=10090;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Kidney;
RA Medline=99113835; PubMed=9916797;
RX Nezu J., Tamai I., Oku A., Ohashi R., Yabuchi H., Hashimoto N., Nikaide H., Sai Y., Koizumi A., Shoji Y., Takada G., Matsushita T., Yashino M., Kato H., Ohura T., Tsujimoto G., Hayakawa J., Shimane M., Tsuji A.;
RA "Primary systemic carnitine deficiency is caused by mutations in a gene encoding sodium ion-dependent carnitine transporter.";
RT Nat. Genet. 21:91-94(1999).
[2]
RN SEQUENCE FROM N.A., AND VARIANT JVS ARG-352.
RP STRAIN=C3H;
RA Medline=99057546; PubMed=9837751;
RX Lu K., Nishimori H., Nakamura Y., Shima K., Kuwajima M.;
RA "A missense mutation of mouse OCTN2, a sodium-dependent carnitine cotransporter, in the juvenile visceral steatosis mouse.";
RT Biochem. Biophys. Res. Commun. 252:590-594(1998).
[3]
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RP MEDLINE=99384224; PubMed=10454528;
RX Chen J., Conway S.J., Ganapathy V.;
RA Wu X., Huang W., Prasad P.D., Seth P., Rajan D.P., Leebach F.H.;
RA "Functional characteristics and tissue distribution pattern of organic cation transporter 2 (OCTN2), an organic cation/carnitine transporter.";
RT J. Pharmacol. Exp. Ther. 290:1482-1492(1999).
[4]
RN FUNCTION: SODIUM-ION DEPENDENT, HIGH AFFINITY CARNITINE TRANSPORTER. ALSO TRANSPORTS ORGANIC CATIONS WITHOUT THE INVOLVEMENT OF SODIUM. INVOLVED IN THE ACTIVE CELLULAR UPTAKE OF CARNITINE.
[5]
RN SUBCELLULAR LOCATION: Integral membrane protein.
[6]
RN DISEASE: DEFECTS IN SLC22A5 ARE THE CAUSE OF JUVENILE VISCERAL STEATOSIS (JVS).
[7]
RN SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. ORGANIC CATION SUBFAMILY.
[8]
RN This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AB015800; BAA36590.1; -
 CC EMBL; AF111425; AAC99787.1; -
 CC EMBL; AF110417; AAD54050.1; -
 CC MGI; MGI:1329012; S1c22a5.
 DR InterPro; IPR003662; sub_transporter.
 DR Pfam; PF00083; sugar_tr.1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 KW Transport; Transmembrane; Glycoprotein; Disease mutation.
 FT TRANSMEM 21 41
 FT TRANSMEM 143 163
 FT TRANSMEM 173 193
 FT TRANSMEM 198 218
 FT TRANSMEM 233 253
 FT TRANSMEM 258 278
 FT TRANSMEM 342 362
 FT TRANSMEM 374 394
 FT TRANSMEM 407 427
 FT TRANSMEM 431 451
 FT TRANSMEM 489 509
 FT CARBOHYD 57 64
 FT CARBOHYD 91 91
 FT CARBOHYD 322 322
 FT CARBOHYD 352 352
 FT VARIANT 352 352 L -> R (IN JVS).
 SQ SEQUENCE 557 AA; 62779 MW; 6093F0EB9612B204 CRC64;

Query Match 15.4%; Score 216.5; DB 1; Length 557;
 Best Local Similarity 23.4%; Pred. No. 1.2e-11;
 Matches 91; Conservative 34; Mismatches 89; Indels 175; Gaps 14;

QY 3 FEELLOVGLGRFOMHLVFLPMLLPH-----ILENFAALPGHRCV-HMLD 55
 DB 4 YDEVTAFLGEMGFQ--RLIFFLSAS-IIPNGFNGMSIV--FLAGTPHRCVPTVN 57
 QY 56 NNTGSGNETGILSEDLRLRISIP---DSNLREKCRFPHPOMQLHNGTSTSEAD 112
 DB 58 LSSAVNNH-----SIPLETKRGQVQKCRRY-----RLATIANSELG 96
 QY 113 TEP-----CVNGWVYDOSYFPTSTYTKMDVCDYOSLSKVQFLLLTGMLVG 159
 DB 97 LEGRDVLDLEQESCLDGMEDYKDVFLSTVTEMDLVCKDKMKAPLITSLFPGVLMG 156
 QY 160 GILGHVSDR----- 169
 DB 157 SPISGQLSDRFGRKNVFLITMGKOTGFSLQVSVNEFEFTVLVGVGQISNVYAFV 216
 QY 170 ----- 216
 DB 217 LGTEILSKSIRITFAITVLCIPYAFGFMYVLPFAFYFIRDMRMLLALVPGVLCALMWF 276
 QY 172 V-ESARMLITINKLDEGLKALKRVARTNGIKNAEETLNEVVRSTM---QEELDAQTKT 227
 DB 277 IPESPRMILISGRIFEAAYIIRKAKINGI-----VAPSTIIFPSELQDINSTRK 326
 QY 228 T---VCDLFNPSMKRKCILVLFKKKIS 253
 DB 327 QLHHITDILFRIRIVITIMSIILMTLITIS 355

RESULT 2
 OCN2_RAT STANDARD; PRT; 557 AA.
 AC 070594; 09QWL0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Organic cation/carnitine transporter 2 (Solute carrier family 22, member 5) (High-affinity sodium-dependent carnitine cotransporter)

DE (US2r) (C11).
 GN SUC22A5 OR OCTN2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA MEDLINE=98200080; PubMed=9541011;
 RA Schoenly E., Spitzberger F., Engelhardt M., Martel F., Oerding N.,
 RA Gruendemann D.;
 RT "Molecular cloning and characterization of two novel transport
 proteins from rat kidney";
 RN FEBS Lett. 425:79-86(1998).
 RP [2]
 RC SEQUENCE FROM N.A.
 RA STRAIN=SPRAGUE-DAWLEY; TISSUE=Intestine;
 RA MEDLINE=99011422; PubMed=9792817;
 RA Sekine T., Kusuhara H., Utsunomiya-Tate N., Tsuda M., Sugiyama Y.,
 RA Kanai Y., Endou H.;
 RT "Molecular cloning and characterization of high-affinity carnitine
 transporter from rat intestine";
 RN Biochem. Biophys. Res. Commun. 251:586-591(1998).
 RP [3]
 RC SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RA MEDLINE=99384224; PubMed=10454528;
 RA Wu X., Huang W., Prasad P.D., Seth P., Rajan D.P., Leibach F.H.,
 RA Chen J., Conway S.J., Ganapathy V.;
 RT "Functional characteristics and tissue distribution pattern of organic
 cation transporter 2 (OCTN2), an organic cation/carnitine
 transporter";
 RN J. Pharmacol. Exp. Ther. 290:1482-1492(1999).
 RC -1- FUNCTION: SODIUM-ION DEPENDENT, HIGH AFFINITY CARNITINE
 TRANSPORTER. ALSO TRANSPORTS ORGANIC CATIONS WITHOUT THE
 INVOLVEMENT OF SODIUM. INVOLVED IN THE ACTIVE CELLULAR UPTAKE OF
 CARNITINE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE PROXIMAL AND DISTAL TUBULES
 AND IN THE GLOMERULI IN THE KIDNEY, IN THE MYOCARDIUM, VALVES, AND
 ARTERIOLES IN THE HEART, IN THE LABYRINTHINE LAYER OF THE
 PLACENTA, AND IN THE CORTEX, HIPPOCAMPUS, AND CEREBELLUM IN THE
 BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. ORGANIC
 CATION SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; AJ001933; CAA05106.1; -
 CC EMBL; AB017260; BAA34399.1; -
 CC EMBL; AF110416; AAD54059.1; -
 CC InterPro; IPR003662; sub_transporter.
 DR Pfam; PF00083; sugar_tr.1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 KW Transport; Transmembrane; Glycoprotein.
 FT TRANSMEM 21 41
 FT TRANSMEM 143 163
 FT TRANSMEM 173 193
 FT TRANSMEM 198 218
 FT TRANSMEM 233 253
 FT TRANSMEM 258 278
 FT TRANSMEM 342 362
 FT TRANSMEM 374 394
 FT TRANSMEM 407 427
 FT TRANSMEM 431 451
 FT TRANSMEM 489 509
 FT CARBOHYD 57 57
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match	Best Local Similarity	Score	DB 1	Length	557
Matches 83; Conservative	77	Mismatches 106;	Indels 151;	Gaps 12;	
3	FEELISGVGGIGRRQMLHVFILPISLMLLPH--ILLNFPAALIGCHRCWAMLDN-NT 58				
4	YDEVTAFLGEMGPQ--RLIFLILSAS-IIPNGFTGLSSVFLATPEHRCRPVDANLSS 60				
59	GSNGETGLISDALLRISIPL--DSNLARECKRRF--VHPQMOLHLN-GTIIHSTSEA 111				
61	AMRNHT-----VPLRLDGRREYVHSCRRRLATIANFSAALGLPEGRDVLGL 108				
112	DTECVGVGVVYDOSYFSTVTKMDVLCVQSLAKSVVQFLLLTGMVGGITGCHVSDR-- 169				
109	EQEESCDGMFESQDVIYIIVTIVEMNLVCEBDMWKAPLTLISLFVGVGLGSGFISGOLSDREG 168				

QY	170	169
Db	169 RKNVLFYTMGOTGFSFLQIFSKNFEMFVLEVLVGMQISNYAFAVLGTIELLKSVRI	228
QY	170	MLV-ESARHLITN 182
Db	229 IFSTGVCIFYAFGYMVLPLFAFYIRDRMLLVALTMPGVCVALMWITPSPKRLISOG	288
QY	183 KIDELKMLRYARANGIKNAEETINIEVST	236
Db	269 RFEELVIRKRAAKNGI- VVSTIFDPSELQDLSKKKQOSHIDLLRTW 338	
QY	237 SMKRRCILVEFLRKKS 253	
Db	339 NIRMVITMSIMLMWTIS 355	
RESULT	4	
QY	GTR2_HUMAN	
ID	GTR2_HUMAN	STANDARD; PRT; 524 AA.
AC	P1168;	
DT	01-JUL-1989 (Rel. 11, Created)	
DT	01-JUL-1989 (Rel. 11, Last sequence update)	
DE	01-MAR-2002 (Rel. 41, Last annotation update)	
DE	Solute carrier family 2, facilitated glucose transporter, member 2	
GN	(Glucose transporter type 2, liver).	
GN	SIC2A2 OR GLUT2.	
OS	Homo sapiens (human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
OC	NCBI_TaxID=9606;	
OC	[1]	
RC	SEQUENCE FROM N.A. AND VARIANT LEU-68.	
RC	TISSUE=Liver, and kidney.	
RX	MEDLINE=88289735; PubMed=3399500;	
RA	Fukunoto H., Seino S., Imura H., Sieno Y., Eddy R.L., Fukushima Y.,	
RA	Byers M.G., Shows T.B., Bell G.I.;	
RT	"Sequence, tissue distribution, and chromosomal localization of mRNA	
RL	encoding a human glucose transporter-like protein.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 85:5434-5438(1988).	
CC	-1- FUNCTION: FACILITATES GLUCOSE TRANSPORT. THIS ISOFORM LIKELY	
CC	MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA	
CC	MEMBRANE OF HEPATOCYTES AND IS RESPONSIBLE FOR UPTAKE OF GLUCOSE	
CC	BY THE BETA CELLS; MAY COMPRISE PART OF THE GLUCOSE-SENSING	
CC	MECHANISM OF THE BETA CELL. MAY ALSO PARTICIPATE WITH THE	
CC	NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSCYTOPLASMIC TRANSPORT OF	
CC	GLUCOSE IN THE SMALL INTESTINE AND KIDNEY.	
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.	
CC	-1- TISSUE SPECIFICITY: LIVER, INSULIN-PRODUCING BETA CELL, SMALL	
CC	INTESTINE, AND KIDNEY.	
CC	-1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE	
CC	TRANSPORTER SUBFAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.1sb-sib.ch/announce/	
CC	or send an email to license@sib-sib.ch)	
CC	-----	
DR	EMBL: J03810; AAA59514.1; -.	
DR	PIR: A31318; A31318.	
DR	MIM: 138160; -.	
DR	InterPro: IPR003663; Sugar transporter.	
DR	InterPro: IPR003662; sub. transporter.	
DR	Pfam: PF00083; sugar_tr.1.	
DR	PRINTS: PRC0171; SUGARTRANSPO.	
DR	PROSITE: PS00216; SUGAR_TRANSPOR_1; 1.	
DR	PROSITE: PS00217; SUGAR_TRANSPOR_2; 1.	
DR	Transmembrane; Sugar transporter; Transport; glycoprotein;	
DR	Multigene family; Polymorphism.	

```

FT DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 11 31 1 (POTENTIAL).
FT DOMAIN 32 98 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 99 119 2 (POTENTIAL).
FT DOMAIN 120 127 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 128 148 3 (POTENTIAL).
FT DOMAIN 149 158 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 159 175 4 (POTENTIAL).
FT DOMAIN 180 208 5 (POTENTIAL).
FT TRANSMEM 209 217 6 (POTENTIAL).
FT DOMAIN 218 238 7 (POTENTIAL).
FT TRANSMEM 239 303 8 (POTENTIAL).
FT TRANSMEM 304 324 9 (POTENTIAL).
FT TRANSMEM 325 359 10 (POTENTIAL).
FT TRANSMEM 339 359 11 (POTENTIAL).
FT TRANSMEM 360 368 12 (POTENTIAL).
FT TRANSMEM 369 389 13 (POTENTIAL).
FT TRANSMEM 390 400 14 (POTENTIAL).
FT TRANSMEM 401 421 15 (POTENTIAL).
FT TRANSMEM 422 433 16 (POTENTIAL).
FT TRANSMEM 434 454 17 (POTENTIAL).
FT TRANSMEM 455 461 18 (POTENTIAL).
FT TRANSMEM 462 482 19 (POTENTIAL).
FT TRANSMEM 483 524 20 (POTENTIAL).
FT TRANSMEM 524 524 21 (POTENTIAL).
FT TRANSMEM 524 524 22 (POTENTIAL).
FT TRANSMEM 524 524 23 (POTENTIAL).
FT TRANSMEM 524 524 24 (POTENTIAL).
FT TRANSMEM 524 524 25 (POTENTIAL).
FT TRANSMEM 524 524 26 (POTENTIAL).
FT TRANSMEM 524 524 27 (POTENTIAL).
FT TRANSMEM 524 524 28 (POTENTIAL).
FT TRANSMEM 524 524 29 (POTENTIAL).
FT TRANSMEM 524 524 30 (POTENTIAL).
FT TRANSMEM 524 524 31 (POTENTIAL).
FT TRANSMEM 524 524 32 (POTENTIAL).
FT TRANSMEM 524 524 33 (POTENTIAL).
FT TRANSMEM 524 524 34 (POTENTIAL).
FT TRANSMEM 524 524 35 (POTENTIAL).
FT TRANSMEM 524 524 36 (POTENTIAL).
FT TRANSMEM 524 524 37 (POTENTIAL).
FT TRANSMEM 524 524 38 (POTENTIAL).
FT TRANSMEM 524 524 39 (POTENTIAL).
FT TRANSMEM 524 524 40 (POTENTIAL).
FT TRANSMEM 524 524 41 (POTENTIAL).
FT TRANSMEM 524 524 42 (POTENTIAL).
FT TRANSMEM 524 524 43 (POTENTIAL).
FT TRANSMEM 524 524 44 (POTENTIAL).
FT TRANSMEM 524 524 45 (POTENTIAL).
FT TRANSMEM 524 524 46 (POTENTIAL).
FT TRANSMEM 524 524 47 (POTENTIAL).
FT TRANSMEM 524 524 48 (POTENTIAL).
FT TRANSMEM 524 524 49 (POTENTIAL).
FT TRANSMEM 524 524 50 (POTENTIAL).
FT TRANSMEM 524 524 51 (POTENTIAL).
FT TRANSMEM 524 524 52 (POTENTIAL).
FT TRANSMEM 524 524 53 (POTENTIAL).
FT TRANSMEM 524 524 54 (POTENTIAL).
FT TRANSMEM 524 524 55 (POTENTIAL).
FT TRANSMEM 524 524 56 (POTENTIAL).
FT TRANSMEM 524 524 57 (POTENTIAL).
FT TRANSMEM 524 524 58 (POTENTIAL).
FT TRANSMEM 524 524 59 (POTENTIAL).
FT TRANSMEM 524 524 60 (POTENTIAL).
FT TRANSMEM 524 524 61 (POTENTIAL).
FT TRANSMEM 524 524 62 (POTENTIAL).
FT TRANSMEM 524 524 63 (POTENTIAL).
FT TRANSMEM 524 524 64 (POTENTIAL).
FT TRANSMEM 524 524 65 (POTENTIAL).
FT TRANSMEM 524 524 66 (POTENTIAL).
FT TRANSMEM 524 524 67 (POTENTIAL).
FT TRANSMEM 524 524 68 (POTENTIAL).
FT TRANSMEM 524 524 69 (POTENTIAL).
FT TRANSMEM 524 524 70 (POTENTIAL).
FT TRANSMEM 524 524 71 (POTENTIAL).
FT TRANSMEM 524 524 72 (POTENTIAL).
FT TRANSMEM 524 524 73 (POTENTIAL).
FT TRANSMEM 524 524 74 (POTENTIAL).
FT TRANSMEM 524 524 75 (POTENTIAL).
FT TRANSMEM 524 524 76 (POTENTIAL).
FT TRANSMEM 524 524 77 (POTENTIAL).
FT TRANSMEM 524 524 78 (POTENTIAL).
FT TRANSMEM 524 524 79 (POTENTIAL).
FT TRANSMEM 524 524 80 (POTENTIAL).
FT TRANSMEM 524 524 81 (POTENTIAL).
FT TRANSMEM 524 524 82 (POTENTIAL).
FT TRANSMEM 524 524 83 (POTENTIAL).
FT TRANSMEM 524 524 84 (POTENTIAL).
FT TRANSMEM 524 524 85 (POTENTIAL).
FT TRANSMEM 524 524 86 (POTENTIAL).
FT TRANSMEM 524 524 87 (POTENTIAL).
FT TRANSMEM 524 524 88 (POTENTIAL).
FT TRANSMEM 524 524 89 (POTENTIAL).
FT TRANSMEM 524 524 90 (POTENTIAL).
FT TRANSMEM 524 524 91 (POTENTIAL).
FT TRANSMEM 524 524 92 (POTENTIAL).
FT TRANSMEM 524 524 93 (POTENTIAL).
FT TRANSMEM 524 524 94 (POTENTIAL).
FT TRANSMEM 524 524 95 (POTENTIAL).
FT TRANSMEM 524 524 96 (POTENTIAL).
FT TRANSMEM 524 524 97 (POTENTIAL).
FT TRANSMEM 524 524 98 (POTENTIAL).
FT TRANSMEM 524 524 99 (POTENTIAL).
FT TRANSMEM 524 524 100 (POTENTIAL).

```

```

RA Amano T., Shibasaki Y., Iin J.L., Akanuma Y., Takaku F., Oka Y.;
RT "The nucleotide sequence of cDNA for a mouse liver-type glucose
RT transporter protein."
RL Nucleic Acids Res. 17:6386-6386(1989).
[3]
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batova S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochenush J.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiraldi L.M., Staehli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlay M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszewski B., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
[4]
RC STRAIN=C57BL/6J;
RX MEDLINE=92111400; PubMed=1165007;
RA Hogan A., Heyner S., Charbon M.J., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A., Thorens B., Schultz G.A.;
RT "Glucose transporter gene expression in early mouse embryos."
RL Development 113:363-372(1991).
[5]
RC STRAIN=C57BL/6J;
RX MEDLINE=93170163; PubMed=1289053;
RA Smith D.E., Gridley T.;
RT "Different screening of a PCR-generated mouse embryo cDNA library:
RT glucose transporters are differentially expressed in early
RT postimplantation mouse embryos."
RL Development 116:555-561(1992).
-1- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM LIKELY
-1- MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
-1- MEMBRANE OF HEPATOCYTES AND IS RESPONSIBLE FOR UPTAKE OF GLUCOSE
-1- BY THE BETA CELLS. MAY COMPRISE PART OF THE GLUCOSE-SENSING
-1- MECHANISM OF THE BETA CELL. MAY ALSO PARTICIPATE WITH THE
-1- NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSCELLULAR TRANSPORT OF
-1- GLUCOSE IN THE SMALL INTESTINE AND KIDNEY.
-1- SUBCELLULAR LOCATION: Integral membrane protein.
-1- TISSUE SPECIFICITY: IN EMBRYO, EXPRESSED IN ENDODERM LAYER OF
-1- YOLK SAC AND LIVER PRIMORDIUM.
-1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
-1- TRANSPORTER SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
-----
DR EMBL: X16986; CAA34855.1; -
DR EMBL: X15684; CAA33719.1; -
DR EMBL: S77926; AAB20847.1; -
DR EMBL: AK005068; BAB23792.1; -
DR PIR: S06920; S06920.
DR PIR: S05319; S05319.
DR MGD: MGI:1095438; Slc2a2.
DR InterPro: IPR003663; Sugar_transporter.

```

DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 DR PRINTS: PR00171; SUGRTNSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane: Sugar transport; Transport; Glycoprotein;
 MW Multigene family.
 FT DOMAIN 1 10
 FT TRANSSEM 11 31 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 32 97 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 98 118 2 (POTENTIAL).
 FT DOMAIN 119 126 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 127 147 3 (POTENTIAL).
 FT DOMAIN 148 157 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 158 178 4 (POTENTIAL).
 FT DOMAIN 179 186 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 187 207 5 (POTENTIAL).
 FT DOMAIN 208 216 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 217 237 6 (POTENTIAL).
 FT DOMAIN 238 302 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 303 323 7 (POTENTIAL).
 FT DOMAIN 324 337 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 338 358 8 (POTENTIAL).
 FT DOMAIN 359 367 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 368 388 9 (POTENTIAL).
 FT DOMAIN 402 422 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 423 432 10 (POTENTIAL).
 FT DOMAIN 433 453 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 454 461 11 (POTENTIAL).
 FT DOMAIN 461 481 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 482 523 12 (POTENTIAL).
 FT DOMAIN 523 523 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 62 62 N-LINKED (GLUCNAC. .) (POTENTIAL).
 FT CONFLICT 106 106 G->D (IN REF. 2).
 FT CONFLICT 181 181 T->N (IN REF. 1).
 FT CONFLICT 327 327 S->T (IN REF. 2).
 FT CONFLICT 351 351 F->L (IN REF. 1).
 FT CONFLICT 432 432 P->S (IN REF. 1).
 SO SEQUENCE 523 AA; 57106 MW; F12B537C0BDB5F63 CRC64;

Query Match Best Local Similarity 6.8%; Score 95; DB 1; Length 523;
 Matches 34; Conservative 27; Mismatches 36; Indels 36; Gaps 6;

OY 143 SLKSVQVFLTLTGMVGGITG-----GHVSDRWLV-----ESARW 177
 DB 186 ALGTLHQLALVTGILISQIAGLSFLLGN-ODWHIILGLSAVPALLQCLLLECESPRY 244
 OY 178 LIITNKLDGLKALKRKVARTNGIKNAETLINIEVVRSTMOEELDAQT-KTYVCDLFRN 235
 DB 245 LVI-KLEEVRAKKSKLRKGTEDYTKDIN-----EMKKEKEASTEOKVSVIQLFTD 296
 OY 236 PSMKRKICLIVFL 248
 DB 297 ANYROPILVALML 309
 RESULT 6
 C11A_MOUSE STANDARD: PRT; 526 AA.
 ID C11A_MOUSE AC 09Q282;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cytochrome P450 11A1, mitochondrial precursor (BC 1.14.15.6) (CYPIA1) (P450(scc)) (Cholesterol side-chain cleavage enzyme) (Cholesterol desmolase).
 DE GN CYP11A OR CYP11A1.
 OC Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RA Tanaka M., Hennebold J.H., Adashi E.Y.;
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE SIDE-CHAIN CLEAVAGE REACTION OF
 CC CHOLESTEROL TO PREGNENOLONE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Cholesterol + reduced adrenal ferredoxin +
 CC O(2) = pregnenolone + 4-methylpentanal + oxidized adrenal
 CC ferredoxin + H(2)O.
 CC -1- PATHWAY: INITIAL RATE-LIMITING REACTION IN THE SYNTHESIS OF
 CC VARIOUS STEROID HORMONES.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF195119; AAF03897.1; -
 CC HSSP: P00189; 1SCC.
 DR MGI:88582; Cyp11a.
 DR InterPro: IPR001128; Cyt_P450.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Membrane; Heme; Mitochondrion;
 KW Steroidogenesis; Transit peptide.
 FT TRANSIT 1 36 MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 37 526 CYTOCHROME P450 11A1.
 FT BINDING 458 458 HEME (BY SIMILARITY).
 SO SEQUENCE 526 AA; 60314 MW; E5029738EEAECB71 CRC64;

Query Match Best Local Similarity 6.7%; Score 93.5; DB 1; Length 526;
 Matches 55; Conservative 38; Mismatches 67; Indels 89; Gaps 13;

OY 51 VHMIDNNTGSGNETGILSEDLRLISIP-----LDSNLRPECRFPVHQWQ 97
 DB 179 LHRRIKQNSGNSGVISD-LRFSPFESSIVFGRMGLEIYVP-AQRFINAVYQ 236
 OY 98 LHLNFTI-----HST-----SEADTEPCVDGMYDQSYFPSTIVT 133
 DB 237 MHTSVPMNLPPDEFRLRTKTKWDAAMADYIFNKAD-----TYQNFY----- 282
 OY 134 KMDL--VCDYQSLKSVVOFL-----LFGMLVGGIIGHVSDRW-LVESARW 177
 DB 283 -WDLRQKRFSQPGVLYSLGSKNKLPEFNQIANITELAGVDTSMTLQMWLYEWAHN 341
 OY 178 LIITNKLDGLKALKRKVARTNGIK-----NAEETLINIEVVRSTMOEELDAQT 227
 DB 342 LKVGEMLRRAVLAARQAGDAKWKQVLPLKASIKETLRHPISVTL-----RY 393
 OY 228 TYVCDL-FRN 235
 DB 394 TVNDLVLRN 402
 RESULT 7
 GTR2_RAT STANDARD: PRT; 522 AA.
 ID GTR2_RAT AC P12336;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Solute carrier family 2, facilitated glucose transporter, member 2
 DE (Glucose transporter type 2, liver).
 GN SLC2A2 OR GLUT2 OR GLUT-2.

Query Match	6.6k	Score 93	DB 1	Length 522
OS Rattus norvegicus (Rat).				
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
RA Thoraxes B., Sarkar H.R., Kaback H.R., Lodish H.F.;				
RT "Cloning and functional expression in bacteria of a novel glucose transporter present in liver, intestine, kidney, and beta-pancreatic islet cells."				
RT Cell 55:261-290(1988).				
RN [1]				
RP SEQUENCE FROM N.A.				
CC MEDLINE=98003066; PubMed=3048704;				
CC -1 FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM LIKELY				
CC MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA				
CC MEMBRANE OF HEPATOCYTES AND IS RESPONSIBLE FOR UPTAKE OF GLUCOSE				
CC BY THE BETA CELLS; MAY COMPRISE PART OF THE GLUCOSE-SENSING				
CC MECHANISM OF THE BETA CELL. MAY ALSO PARTICIPATE WITH THE				
CC NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSCYLLULAR TRANSPORT OF				
CC GLUCOSE IN THE SMALL INTESTINE AND KIDNEY.				
CC -1 SUBCELLULAR LOCATION: Integral membrane protein.				
CC -1 TISSUE SPECIFICITY: PRESENT IN LIVER, INTESTINE, KIDNEY AND				
CC BETA-PANCREATIC ISLET CELLS.				
CC -1 SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE				
CC TRANSPORTER SUBFAMILY.				
CC -----				
CC This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC the European Bioinformatics Institute. There are no restrictions on its				
CC use by non-profit institutions as long as its content is in no way				
CC modified and this statement is not removed. Usage by and for commercial				
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC or send an email to license@isb-sib.ch).				
CC -----				
CC EMBL, J03145; AAA41298.1; -				
DR PIR, A31556; A31556.				
DR InterPro: IPR003663; Sugar_transporter.				
DR InterPro: IPR003662; sub-transporter.				
DR Pfam: PF00083; sugar_tr. 1.				
DR PRINTS: PR00171; SUGTRNSPORT.				
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.				
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.				
KW Transmembrane; Sugar transport; transporter; glycoprotein;				
KW Multigene family.				
FT DOMAIN 1 10				
FT TRANSMEM 1 1 31				
FT DOMAIN 32 96				
FT TRANSMEM 97 117				
FT DOMAIN 118 125				
FT TRANSMEM 126 146				
FT DOMAIN 147 156				
FT TRANSMEM 157 177				
FT DOMAIN 178 185				
FT TRANSMEM 186 206				
FT DOMAIN 207 215				
FT TRANSMEM 216 236				
FT DOMAIN 237 301				
FT TRANSMEM 302 322				
FT DOMAIN 323 337				
FT TRANSMEM 338 358				
FT DOMAIN 359 365				
FT TRANSMEM 366 386				
FT DOMAIN 387 401				
FT TRANSMEM 402 422				
FT DOMAIN 423 431				
FT TRANSMEM 432 452				
FT DOMAIN 453 459				
FT TRANSMEM 460 480				
FT DOMAIN 481 522				
FT TRANSMEM 522 522				
FT CARBOHYD N-LINKED (GLCNAC. .) (POTENTIAL).				
FT SEQUENCE 522 AA; 57085 MW; 075AB81B56CF33F7 CRC64;				

```
Best Local Similarity 24.8%; Pred.No.0.97; Matches 33; Conservative 28; Mismatches 36; Indels 36; Gaps 6;
```

```
OY      143 SLKSVQFLLLTGMVGLIG-----GHVSDFLV-----ESARW 177  
          :|::||::|: | | | | | | | | | | | | | | | |  
Db       185 ALGTHLAVLTGIIISQIAGLSFIIGN-ODYWHLILGISAPALLIQCLILLFCPESPRY 243  
  
OY      178 LIINIKDEGIKAIRKVARTNGIKNAEETLIIVNVSTMOEELDAQT--KTVCDFLRN 235  
          ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db       244 LYL--NLSEEVRAKSLKLRLGTEDTIKDIN-----EMRKEKEASTEGKVASVIQLTFMD 295  
  
OY      236 PSMRRICITVELF 248  
          |:|:|:|:|:  
Db       296 PNRYOPIVVALML 308
```

```
RESULT      8  
FKBX_MOUSE  STANDARD;              PRI:    561 AA.  
AC           061576;  
DT     30-MAY-2000 (Rel. 39, Created)  
DT     30-MAY-2000 (Rel. 39, Last sequence update)  
DT     16-OCT-2001 (Rel. 40, Last annotation update)  
DE     65 kDa FKBP6-binding protein precursor (EC 5.2.1.8) (FKBP65) (FKBPRP)  
DE     (peptidyl-L-prolyl cis-trans isomerase) (Piase) (Rotamase)  
DE     (Immunophilin FKBP65).  
GN     FKBP6 OR FKBP1-RS OR FKBP-RS OR FKBPRP.  
OS     Mus musculus (Mouse).  
OC     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX     NCBI_TaxId=10090;  
RN       [1]  
RP       SEQUENCE FROM N.A.  
RC     STRAIN-JB6; TISSUE-Epidermis;  
RX     MEDLINE=94117013; PubMed=7507077;  
RA     Simek S.L., Kozak C.A., Winterstein D., Hegamyer G., Colburn N.H.;  
RT     "Sequence and localization of a novel FKBP5-binding protein to mouse  
        chromosome 11.";  
RT     Genomics 18:407-409(1993).  
RL       [2]  
RN       SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RP       MEDLINE=96094328; PubMed=7493967;  
RX     Cross M.C., Winterstein D., Sowder R.C. II, Simek S.L.;  
RA     "Molecular cloning, DNA sequence analysis, and biochemical  
        characterization of a novel 65-kDa FKBP5-binding protein (FKBP65).";  
RT     J. Biol. Chem. 270:29336-29341(1995).  
CC       -1- FUNCTION: PIASES ACCELERATE THE FOLDING OF PROTEINS DURING  
         PROTEIN SYNTHESIS.  
CC       -1- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC  
         PEPTIDE BONDS IN OLIGOPEPTIDES.  
CC       -1- ENZYME REGULATION: INHIBITED BY BOTH FKBP56 AND RAPAMYCIN, BUT NOT  
         BY CYCLOSPORINE A.  
CC       -1- SUBCELLULAR LOCATION: ENDOLYSAMIC RETICULUM (POTENTIAL).  
CC       -1- TISSUE SPECIFICITY: EXPRESSED IN THE LUNG, SPLEEN, HEART, BRAIN  
         AND TESTIS.  
CC       -1- PTM: GLYCOSYLATED AND PHOSPHORYLATED.  
CC       -1- SIMILARITY: BELONGS TO THE FKBP-TYPE PIASE FAMILY. CONTAINS 4  
         FKBP-LIKE DOMAINS.  
CC       -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
```

```
--  
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation --  
CC   the European Bioinformatics institute. There are no restrictions on how  
CC   use by non-profit institutions as long as its content is in no way  
CC   modified and this statement is not removed. Usage by and for commercial  
CC   entities requires a license agreement (see http://www.isb-slb.ch/announce/  
CC   or send an email to licenses@slb-slb.ch).
```

```
--  
DR     EMBL; I07063; AAC37678.1; .  
DR     HSSP; P20081; IYAT.  
DR     MGDI; I04769; FKBP6.  
DR     InterPro; IPRO02048; EF-hand.  
DR     InterPro; IPRO00886; ER-target.
```

DR InterPro: IPR001179; FKBP_PPIase.
 DR Pfam: PF00036; ehand: 2.
 DR Pfam: PF00254; FKBP: 4.
 DR SMART: SM00054; EFh: 2.
 DR PROSITE: PS00453; FKBP_PPIASE_1; FALSE_NEG.
 DR PROSITE: PS00454; FKBP_PPIASE_2; 1.
 DR PROSITE: PS00059; FKBP_PPIASE_3; 4.
 DR PROSITE: PS00114; ER_TARGET; 1.
 DR PROSITE: PS00018; EF_HAND; 1.
 DR Isomerase; Retanase; Repeat; Glycoprotein; Signal; Phosphorylation;
 KM Endoplasmic reticulum; Calcium-binding.
 FT SIGNAL 1 33
 FT CHAIN 34 581
 FT DOMAIN 54 146
 FT DOMAIN 179 258
 FT DOMAIN 291 370
 FT DOMAIN 389 482
 FT CA_BIND 509 520
 FT CA_BIND 554 565
 FT CARBOHYD 69 69
 FT CARBOHYD 181 181
 FT CARBOHYD 293 309
 FT CARBOHYD 351 351
 FT CARBOHYD 392 392
 FT CARBOHYD 406 406
 FT SITE 578 581
 FT SEQUENCE 581 AA; 64669 MW; 1B51B3032089F555 CRC64;

Query Match 6.2%; Score 87; DB 1; Length 581;
 Best Local Similarity 25.1%; Pred. No. 3.8;
 Matches 53; Conservative 24; Mismatches 66; Indels 68; Gaps 10;

QY 19 LHVFLPSMLLPHLLLENFAAIPGHCWVHMLDNTSGNETGILSEDAAL-RISI 77
 DB 11 LHVRLILPLLL-----QLEGLGRASAGAPLDEVIYERHI 50
 QY 78 PLDSNLRPEKCRFVH-PQMLHLNCTIHTSTADTEPCVDGVYDOSYSPSTIVTKMD 136
 DB 51 -----PRACREPVOMGDFVRYHNGTFE-----DGKKFPSSYDRSLV--- 88
 QY 137 LVCDYSLKSVQFL--LLTGMVGGIIGHVSDFRMLVESARWLIITNKLDGKALKRV 194
 DB 89 -----AIVGVGRLLTGM-DRGLMGCVNER-----RLIVPHLGYSGIVA-- 130
 QY 195 ARTNGIKNAEETINIEVVRSTMOEELDAQT 225
 DB 131 ----GLIPDAILYFDVLDVWVKADTVQS 157

RESULT 9
 GTR3_CHICK STANDARD; PRT; 496 AA.
 AC P28568;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Solute carrier family 2, facilitated glucose transporter, member 3
 DE (Glucose transporter type 3) (CEF-GT3).
 GN SLC2A3 OR GLUT3.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91342646; PubMed=1875932;
 RA White M.K., Rall T.B., Weber M.J.;
 RT "Differential regulation of glucose transporter isoforms by the src
 oncogene in chicken embryo fibroblasts."; Mol. Cell. Biol. 11:4448-4454(1991).

CC -1- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. PROBABLY A NEURONAL
 CC GLUCOSE TRANSPORTER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
 CC TRANSPORTER SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M37785; AAA48662.1; -
 CC PIR: A41264; A41264.
 DR InterPro: IPR003663; Sugar_transporter.
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 DR PRINTS: PR00171; SUGAR_TRANSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane; Sugar transporter; Transport; Glycoprotein;
 KW Multigene family.
 FT DOMAIN 1 6
 FT TRANSMEM 7 27
 FT TRANSMEM 28 65
 FT TRANSMEM 66 86
 FT DOMAIN 87 94
 FT TRANSMEM 95 115
 FT DOMAIN 116 125
 FT TRANSMEM 126 146
 FT DOMAIN 147 154
 FT TRANSMEM 155 175
 FT DOMAIN 176 184
 FT TRANSMEM 185 205
 FT DOMAIN 206 270
 FT TRANSMEM 271 291
 FT DOMAIN 292 306
 FT TRANSMEM 307 327
 FT DOMAIN 328 335
 FT TRANSMEM 336 356
 FT DOMAIN 357 367
 FT TRANSMEM 368 388
 FT DOMAIN 389 398
 FT TRANSMEM 399 419
 FT DOMAIN 420 426
 FT TRANSMEM 427 447
 FT DOMAIN 448 496
 FT SITE 278 280
 FT CARBOHYD 44 44
 FT SEQUENCE 496 AA; 54174 MW; 75B3C0F61A7A92A5 CRC64;
 N-LINKED (GLCNAC...) (POTENTIAL).
 SIMILARITY).

Query Match 6.2%; Score 86.5; DB 1; Length 496;
 Best Local Similarity 26.5%; Pred. No. 3.4;
 Matches 31; Conservative 26; Mismatches 35; Indels 25; Gaps 5;

QY 135 WDLVCDYSLKSVQFLLTGMVGGIIGHVSDFRMLVESARWLIITNKLDGKALKRV 194
 DB 185 WPLLLGFTIVAVLQVALL-----FCPSPEFLI-NKEBE-RAQTVL 227
 QY 195 ARTNGIKNAEETINIEVVRSTMOEELD--AAQTKTVCDLFENPGRKRICLVFAR 249
 DB 228 QKLRGTQDVSDI-----SEMKESAKMSQEKATVLELFRSPYRPIITISITLQ 278

RESULT 10
 GYR3_THEMA STANDARD; PRT; 636 AA.
 ID GYR3_THEMA
 AC P77993;
 DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA gyrase subunit B (EC 5.99.1.3) (Topoisomerase II subunit B).
 GN GYB OR TOP2B OR TM0833.
 OS Thermotoga maritima.
 CC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_Taxid:2336;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-MSB8 / DSM 3109;
 RA MEDLINE:9701713; PubMed-8863738;
 RA Guipaud O., Labedan B., Forterie P.;
 RT "A gybB-like gene from the hyperthermophilic bacterium Thermotoga
 RT maritima."
 RL Gene 174:121-128(1996).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN-MSB8 / DSM 3109;
 RC MEDLINE:99287316; PubMed-10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 RA Hatt D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of Thermotoga maritima."
 RL Nature 399:323-329(1999).
 CC -1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
 CC STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE
 CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
 CC DNA RINGS. INCLUDING CATENANES AND KNOTTED RINGS.
 CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
 CC of double-stranded DNA.
 CC -1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
 CC BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
 CC ENZYME FORMS AN A2B2 TETRAMER.
 CC -1- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U96922; AAC44498.1; -.
 CC EMBL: AE001750; AAD35915.1; -.
 CC HSSP: P06982; IAJ6.
 CC TIGR: TM0833; -.
 DR InterPro: IPR002288; DNA_gyraseB_C.
 DR InterPro: IPR001241; DNA_topoisomII.
 DR InterPro: IPR003594; HATPase_C.
 DR InterPro: IPR004359; HIS_KIN_ssg.
 DR InterPro: IPR002936; Toprim.
 DR Pfam: PR00986; DNA_gyraseB_C; 1.
 DR Pfam: PR00204; DNA_topoisomII; 1.
 DR Pfam: PR02518; HATPase_C; 1.
 DR Pfam: PF01751; Toprim; 1.
 DR PRINTS: PR00418; TP12FAMILY.
 DR PRODOM: PD000616; DNA_topoisomII; 1.
 DR PRODOM: PD149633; DNA_gyraseB_C; 1.
 DR SMART: SM00387; HATPase_C; 1.
 DR SMART: SM00433; TOP2C; 1.
 DR PROSITE: PS00177; TOPOISOMERASE_II; 1.
 KW Isoenzyme; Topoisomerase; ATP-binding; Complete proteome.
 FT CONFLICT 152 152 V -> M (IN REF. 1).
 SQ SEQUENCE 636 AA; 72427 MW; B21E0ECBC6F89E CRC64;

Matches	57;	Conservative	44;	Mismatches	89;	Indels	95;	Gaps	13;
QY	10	VGGLGRFQMLATVFLIPSMILIPHILLLENFAAATGHCWVHMNDNNGSNGEGLISE	69						
Db	26	IGSTGKRGKIHLLVY-----EVDVNSVDALAGYCDWIKV-----TLHE	63						
QY	70	DALLRI-----SPIDLSNLTPEKCRFPVHPOMOLHLNG-----TI	105						
Db	64	DGSVEYEDNKGRIPLVD--IHPEGRGRALLEVFLVLAGGKFSKDSYKISGLHGVSVV	121						
QY	106	HSTSE-ADTEPCVDGWYVDGSIYFPSTIVTKMDLVCDYOSLSKYQF-----	150						
Db	122	NALSEMLEVHRHVRGKLYRQRYERKGVPTVEVEYIGETDKHGTIVRKPDPLFSETEPD	181						
QY	151	-----LLTGMLVGGI-----IGHVS-----DRMLVESANMLITTKDLGKALK	193						
Db	182	DIEHRLREIAFLVPGKIEFEDRINKETKFEKDGIVAYVYV-----LNGKRALHD	235						
QY	194	V--ARTNGI--KNAEETLNIYV---VNSTOGEELDAQTKTV	229						
Db	236	VIHKRTKVKTKNGEDEVIYELAFQYDYSIEDIVSFANTIKTV	280						
RESULT	11								
ID	GT3-_BOVIN	STANDARD;	PRT;	494	AA.				
AC	P58352;								
DT	01-MAR-2002 (Rel. 41, Created)								
DT	01-MAR-2002 (Rel. 41, Last sequence update)								
DT	01-MAR-2002 (Rel. 41, Last annotation update)								
DE	Solute carrier family 2, facilitated glucose transporter, member 3								
DE	(Glucose transporter type 3, brain).								
GN	SLC2A3 OR GLUT3.								
OS	Bos taurus (Bovine).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;								
OC	Bovidae; Bovinae; Bos.								
OX	NCBI_Taxid:9913;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RA	Augustin R., Fischer B.;								
RT	"Glucose transporter expression during bovine preimplantation embryo								
RT	development."								
RL	Submitted (May-2001) to the EMBL/Genbank/DBJ databases.								
RN	[2]								
RP	SEQUENCE OF 385-494 FROM N.A.								
RC	TISSUE=Brain;								
RA	Augustin R., Navarrete-Santos A., Fischer B.;								
RT	"Expression of glucose transporters amongst ruminants."								
RT	Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.								
RL	-!- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. PROBABLY A NEURONAL								
CC	GLUCOSE TRANSPORTER.								
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.								
CC	-!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE								
CC	TRANSPORTER SUBFAMILY.								
CC									
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/								
CC	or send an email to license@isb-sib.ch).								
CC									
DR	EMBL: AY033938; AAK70222.1; -								
DR	EMBL: AF308829; AAK63202.1; -								
DR	PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.								
DR	PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.								
KW	Transmembrane; Sugar transport; Transport; Glycoprotein;								
KW	Multigene family.								
FT	DOMAIN	1	5						
FT	TRANSMEM	6	26						
FT	DOMAIN	1	26						
FT	DOMAIN	27	64						
FT	DOMAIN	27	64						
FT	DOMAIN	27	64						

```

FT TRANSMEM 65 85 2 (POTENTIAL).
FT DOMAIN 86 93 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 94 114 3 (POTENTIAL).
FT DOMAIN 115 124 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 125 145 4 (POTENTIAL).
FT DOMAIN 146 159 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 160 180 5 (POTENTIAL).
FT DOMAIN 181 183 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 184 204 6 (POTENTIAL).
FT DOMAIN 205 269 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 270 290 7 (POTENTIAL).
FT DOMAIN 291 305 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 306 326 8 (POTENTIAL).
FT DOMAIN 327 334 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 335 355 9 (POTENTIAL).
FT DOMAIN 356 384 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 385 399 10 (POTENTIAL).
FT DOMAIN 400 420 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 421 424 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 425 445 12 (POTENTIAL).
FT TRANSMEM 446 494 CYTOPLASMIC (POTENTIAL).
FT SITE 277 279 DEFINES SUBSTRATE SPECIFICITY (BY SIMILARITY).
FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 358 358 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 385 389 PWETV -> SLVYC (IN REF. 2).
SQ SEQUENCE 494 AA; 34019 MW; EC42A3C648CAD23C CRC64;

```

```

Query Match 5.9%; Score 83.5; DB 1; Length 494;
Best Local Similarity 23.3%; Pred. No. 6.3;
Matches 27; Conservative 26; Mismatches 40; Indels 23; Gaps 3;

```

```

QY 135 WDVLCVQSLKSVQVFLITGMLVGIIGHVSDRWLVESARMLITNKLDEGLK-ALRK 193
DB 184 WPLLGLFTLLPALIOCALP-----FCPESPRFLINRKEEKAKELIOR 228
QY 194 VARTNGIKNAEETLINIEVYSTMOEELDAOQTKTYCDLFNPSMRKRICILVFLR 249
DB 229 LWGTEDVAODIOEKDESMRMSOEKOV-----TVLELFRAPNYROPITISIMIQ 277

```

```

RESULT 12
GTR3_SHEEP STANDARD; PRT; 494 AA.
AC P47843;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 3
DE (Glucose transporter type 3, brain).
DE SLC2A3 OR GLUT3 OR GLUT-3.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COOPMORTH;
RX MEDLINE=96109471; PubMed=8653093;
RA Bennett B.L., Prosser C.G., Grigor M.R.;
RT "Isolation of cDNAs and tissue specific expression of ovine glucose
RT transporters.";
RL Biochem. Mol. Biol. Int. 37:9-16(1995).
RN [2]
RP SEQUENCE OF 88-248 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=97392487; PubMed=9250701;
RA Currie M.J., Bassett N.S., Gluckman P.D.;
RT "Ovine glucose transporter-1 and -3: cDNA partial sequences and
RT developmental gene expression in the placenta.";

```

```

RL Placenta 18:393-401(1997).
CC -1- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. PROBABLY A NEURONAL
CC GLUCOSE TRANSPORTER.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
CC TRANSPORTER SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L39214; AAC41629.1; -.
DR EMBL; U89030; AAB49313.1; -.
DR InterPro; IPR003663; Sugar_transporter.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00063; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Transport; Glycoprotein;
KW Multigene family.
FT DOMAIN 1 5 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 6 26 1 (POTENTIAL).
FT DOMAIN 27 64 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 65 85 2 (POTENTIAL).
FT DOMAIN 86 93 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 94 114 3 (POTENTIAL).
FT DOMAIN 115 124 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 125 145 4 (POTENTIAL).
FT DOMAIN 146 153 4 (POTENTIAL).
FT TRANSMEM 154 174 5 (POTENTIAL).
FT DOMAIN 175 183 5 (POTENTIAL).
FT TRANSMEM 184 204 6 (POTENTIAL).
FT DOMAIN 205 269 6 (POTENTIAL).
FT TRANSMEM 270 290 7 (POTENTIAL).
FT DOMAIN 291 304 7 (POTENTIAL).
FT TRANSMEM 305 325 8 (POTENTIAL).
FT DOMAIN 326 334 8 (POTENTIAL).
FT TRANSMEM 335 355 9 (POTENTIAL).
FT DOMAIN 356 384 9 (POTENTIAL).
FT TRANSMEM 385 399 10 (POTENTIAL).
FT DOMAIN 400 420 10 (POTENTIAL).
FT TRANSMEM 421 424 11 (POTENTIAL).
FT DOMAIN 425 445 11 (POTENTIAL).
FT TRANSMEM 446 494 12 (POTENTIAL).
FT SITE 277 279 DEFINES SUBSTRATE SPECIFICITY (BY SIMILARITY).
FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 494 AA; 54194 MW; AB9204D3EAF4BFBA CRC64;

```

```

Query Match 5.9%; Score 83.5; DB 1; Length 494;
Best Local Similarity 23.3%; Pred. No. 6.3;
Matches 27; Conservative 26; Mismatches 40; Indels 23; Gaps 3;

```

```

QY 135 WDVLCVQSLKSVQVFLITGMLVGIIGHVSDRWLVESARMLITNKLDEGLK-ALRK 193
DB 184 WPLLGLFTLLPALIOCALP-----FCPESPRFLINRKEEKAKELIOR 228
QY 194 VARTNGIKNAEETLINIEVYSTMOEELDAOQTKTYCDLFNPSMRKRICILVFLR 249
DB 229 LWGTEDVAODIOEKDESMRMSOEKOV-----TVLELFRAPNYROPITISIMIQ 277

```

```

RESULT 13
C11A_CAPHI STANDARD; PRT; 520 AA.
ID C11A_CAPHI

```

```

AC      P79153;
DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      15-JUL-1999 (Rel. 38, Last annotation update)
DE      Cytochrome P450 11A1, mitochondrial precursor (EC 1.14.15.6) (CYP11A1)
DE      (P450(scsc)) (Cholesterol side-chain cleavage enzyme) (Cholesterol
DE      desmolase).
GN      CYP11A1
OS      Capra hircus (Goat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC      Bovidae; Caprinae; Capra.
OX      NCBI_TaxID=9925;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Adrenal cortex;
RX      MEDLINE=96244421; PubMed=8645627;
RA      Okuyama E., Okazaki T., Furukawa A., Wu R.-F., Ichikawa Y.;
RT      Molecular cloning and nucleotide sequences of cDNA clones of sheep
RT      and goat adrenocortical cytochromes P450sc (CYP11A1).";
RL      J. Steroid Biochem. Mol. Biol. 57:179-185(1996).
CC      -1- FUNCTION: CATALYZES THE SIDE-CHAIN CLEAVAGE REACTION OF
CC      CHOLESTEROL TO PRENENEOLONE.
CC      -1- CATALYTIC ACTIVITY: Cholesterol + reduced adrenal ferredoxin +
CC      O(2) = pregnenolone + 4-methylpentanal + oxidized adrenal
CC      ferredoxin + H(2)O.
CC      -1- PATHWAY: INITIAL RATE-LIMITING REACTION IN THE SYNTHESIS OF
CC      VARIOUS STEROID HORMONES.
CC      -1- SUBCELLULAR LOCATION: Mitochondrial.
CC      -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; D50058; BAA08776.1; -.
DR      HSSP; P00189; ISCC.
DR      InterPro; IPR001128; CYL_P450.
DR      Pfam; PF00067; P450.1.
DR      PRINTS; PS00385; P450.
DR      PROSITE; PS00086; CYTOCHROME_P450.1.
KW      Oxidoreductase; Monooxygenase; Membrane; Heme; Mitochondrion;
KW      Steroidogenesis; Transil peptide.
KW      TRANSIT 1 39 MITOCHONDRION (BY SIMILARITY).
FT      CHAIN 40 520 CYTOCHROME P450 11A1.
FT      BINDING 461 461 HEME (BY SIMILARITY).
SQ      SEQUENCE 520 AA; 60418 MW; 4FB09A3C89310317 CRC64;

```

Query Match	5.9%	Score 83.5	DB 1	Length 520
Best Local Similarity	25.2%	Pred. No. 6.7		
Matches	51	Conservative	32	Mismatches 90; Indels 29; Gaps 9;
QY	34	HLLENFAALIGHHCWAMLDNNNGSGNGLISEDALRLSRIPDSLRREKORRFVH	93	
DB	204	HAPEFISITNVMEGR--IGMEDIVNTEKQKTDVYKKFHTSVPL-LVPEPLYRLFTT	260	
QY	94	POMOLHLNG--TISTSEADREPCVDGWVYQSYFSPSTIVTKMDLVCDYQSIKSVQEL	151	
DB	261	KWR-DHVAAMPTLINFNAKEYTE-----IFYODLRQKTEFENYPGIL-YHLTKSEKML	312	
QY	152	-----LITGMLVGITIGVSDRW-LVESARMLITNKIDEGKMLRKVARTNGIR----	201	
DB	313	EDVKANITMLAGVDYTSITLQMHLYEMARSLNQEMLRREVLNARRQAEDISKMLQ	372	
QY	202	-----NAEETLNIEVVRSTMQ	217	
DB	373	VPLTKASIKETFLRHPISTLQ	394	

	RESULT	14
CC	GTRL_CHKCK	ID
AC	P46896;	STANDARD;
DT	01-NOV-1995 (Rel. 32, Created)	PRT; 490 AA.
DT	01-MAY-1995 (Rel. 32, Last sequence update)	
DT	01-MAR-2002 (Rel. 41, Last annotation update)	
DE	Solute carrier family 2, facilitated glucose transporter, member 1	
DE	(glucose transporter type 1) (GTL1).	
CN	SIC2A1 OR GLUT1.	
OS	Gallus gallus (chicken).	
OC	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;	
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;	
OC	Gallus.	
OX	NCBI_TaxID=9031;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=96157892; PubMed=8589457;	
RA	Wagstaff P., Kang H.Y., Mylott D., Robbins P.J., White M.K.:	
RT	"Characterization of the avian GLUT1 glucose transporter-	
RT	differential regulation of GLUT1 and GLUT3 in chicken embryo	
RT	fibroblasts."	
RL	Mol. Biol. Cell 6:1575-1589(1995).	
CC	-I- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM MAY BE	
CC	RESPONSIBLE FOR CONSTITUTIVE OR BASAL GLUCOSE UPTAKE. HAS A VERY	
CC	BROAD SUBSTRATE SPECIFICITY; CAN TRANSPORT A WIDE RANGE OF ALDOSES	
CC	INCLUDING BOTH PENTOSES AND HEXOSES (BY SIMILARITY).	
CC	-I- SUBCELLULAR LOCATION: Integral membrane protein. Localizes	
CC	primarily at the cell surface (by similarity).	
CC	-I- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE	
CC	TRANSPORTER SUBFAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; L07300; AB02037.1; -	
DR	InterPro; IPR003663; Sugar_transporter.	
DR	InterPro; IPR003662; sub_transporter.	
DR	Pfam; PF000083; sugar_tr. 1.	
DR	PRINTS; PR00171; SUGETRNSPORT.	
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.	
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.	
KW	Transmembrane; Sugar transport; Glycoprotein;	
KW	Multigene family.	
FT	DOMAIN 1 11	
FT	TRANSMEM 12 32	1 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 33 65	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 66 86	2 (POTENTIAL).
FT	DOMAIN 87 94	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 95 115	3 (POTENTIAL).
FT	DOMAIN 116 125	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 126 146	4 (POTENTIAL).
FT	DOMAIN 147 154	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 155 175	5 (POTENTIAL).
FT	DOMAIN 176 184	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 185 205	6 (POTENTIAL).
FT	DOMAIN 206 270	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 271 291	7 (POTENTIAL).
FT	DOMAIN 292 306	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 307 327	8 (POTENTIAL).
FT	DOMAIN 328 336	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 337 357	9 (POTENTIAL).
FT	DOMAIN 358 369	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 370 390	10 (POTENTIAL).
FT	DOMAIN 391 400	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 401 421	11 (POTENTIAL).
FT	DOMAIN 422 428	EXTRACELLULAR (POTENTIAL).

```

FT TRANSMEM 429 449 12 (POTENTIAL).
FT DOMAIN 450 490 CYTOPLASMIC (POTENTIAL).
FT SITE 177 179 DEFINES SUBSTRATE SPECIFICITY (BY
FT CARBOHYD 44 44 SIMILARITY).
FT SEQUENCE 490 AA: 54086 MW: 638B176812D06104 CRC64;

Query Match
Best Local Similarity 23.3%; Score 82.5; DB 1; Length 490;
Matches 27; Conservative 28; Mismatches 38; Indels 23; Gaps 4;

QY 135 WLVCDYOSLKVYVFLITGMLVGGIIGHVSRLVLESARLLIT-NKIDEGIKALRK 193
DQ 185 WFLGGEFVFPALLOCIILP-----FAPESPRLILNRNEENKAKSVLKK 229
QY 194 VARTNGIKNAEETLINEVVRSTMOEELDAQRTTYCDDLRNFSMRKRICILVFLR 249
DQ 230 LRGITDVSSDLQEMK-EESRQMKREK-----KVTIMELFRSPMYRPIILAIVLQ 278

RESULT 15
GHT3_SCHPO STANDARD; PRI: 555 AA.
ID GHT3_SCHPO
AC 092339;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE High-affinity glucanase transporter ght3 (Hexose transporter 3).
GN GHT3 OR SPAC18.01.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=20200352; PubMed=10735857;
RA Heland S., Radovanovic N., Hoefler M., Winderlich J., Lichtenberg H.;
RT "Multiple hexose transporters of Schizosaccharomyces pombe.";
RL J. Bacteriol. 182:2153-2162(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: HIGH-AFFINITY GLUCONATE TRANSPORTER.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: AF051139; AAC63975.1; -.
CC DR EMBL: 281312; CAB03595.1; -.
CC DR InterPro: IPR003663; Sugar_transporter.
CC DR InterPro: IPR003662; sub_transporter.
CC DR Pfam: PF00063; sugar_tr; 1.
CC DR PRINTS: PR00171; SUGARTRNSPRT.
CC DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
CC DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
CC KW Repeat: Transmembrane; Sugar transporter; Glycoprotein.
CC FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 10 30 1 (POTENTIAL).
CC FT DOMAIN 31 58 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 59 79 2 (POTENTIAL).
CC FT DOMAIN 80 87 CYTOPLASMIC (POTENTIAL).

```

```

FT TRANSMEM 88 108 3 (POTENTIAL).
FT DOMAIN 109 112 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 113 133 4 (POTENTIAL).
FT DOMAIN 134 144 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 145 165 5 (POTENTIAL).
FT DOMAIN 166 179 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 180 200 6 (POTENTIAL).
FT DOMAIN 201 266 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 267 285 7 (POTENTIAL).
FT DOMAIN 286 301 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 302 322 8 (POTENTIAL).
FT DOMAIN 323 328 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 329 349 9 (POTENTIAL).
FT DOMAIN 350 363 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 364 384 10 (POTENTIAL).
FT DOMAIN 385 404 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 405 425 11 (POTENTIAL).
FT DOMAIN 426 432 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 433 453 12 (POTENTIAL).
FT DOMAIN 454 555 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 523 555 SER-RICH.
FT SEQUENCE 555 AA: 62094 MW: 30DF04294D318DB CRC64;

Query Match
Best Local Similarity 5.9%; Score 82.5; DB 1; Length 555;
Matches 52; Conservative 39; Mismatches 91; Indels 51; Gaps 11;

QY 22 VFILPSMLLIPHILLENFAAIPGRWCVMHLDNNTSGNETGISDALRISIPIDS 81
DQ 95 VYIIEELLV-----TAVPS--WIGLVKILAGVIGALS-----LSPGYOS 136
QY 82 NLREKCRREYHPQWQLHLNGTISTSEADTEPCVDGMYDYQSTFPTITVKMDLVCDY 141
DQ 137 EVAPPOIRGAVVATYQ-----IFSTGALVAACINNGTHTLRKTAS----NRTSFGI 184
QY 142 QSLKSVYVFLITGMLVGGIIGHVSRLVLESARLLITNKLDEGLAKRKVARTNGIK 201
DQ 185 NMLMGI---LLMVGVL-----FLPESPRYLILYKGRDEALRLMCMMAELSPSS 229
QY 202 NAEETLINEVVRSTMOEELDAQRTTYCDDLRNFSMRKRICI-LVFLRKKI 252
DQ 230 EIIOT-NENTIKSDI--EIEWAGKARWIEIF-GMDIRYRFLGFLVLMFREL 278

Search completed: July 1, 2002, 11:43:04
Job time: 214 sec

```

Tue Jul 2 10:13:42 2002

us-09-674-235-1.rpr

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

July 1, 2002, 11:37:39 ; Search time 17.64 Seconds
(without alignments)
1459.861 Million cell updates/sec

US-09-674-235-1
1406
1 MAFELLSQVGLGRFOMLH.....RKISRKRHNDCYKTKR 268

Title:
Perfect score:
Sequence:

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR:1:1
2: PIR:2:1
3: PIR:3:1
4: PIR:4:1

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	197.5	14.0	576	T22509	hypothetical prote
2	193.5	13.8	577	JM0089	organic cation tra
3	193	13.7	577	JM0346	high-affinity can
4	169.5	12.1	556	S50862	organic cation tra
5	162.5	11.6	593	U04884	hypothetical prote
6	156	11.1	794	T27870	sugar transport pr
7	108	7.7	591	T30895	probable protein
8	106	7.5	539	C96758	hypothetical prote
9	103	7.3	745	T16565	glucose transport
10	100	7.1	524	A13138	protein ZK455.8
11	98.5	7.0	447	D89646	L-arabinose transp
12	97	6.9	464	F69587	glucose transport
13	95	6.8	523	S06920	probable membrane
14	93	6.6	521	G84864	glucose transport
15	93	6.6	522	A31556	conserved hypothet
16	91	6.5	434	C83395	probable oligopept
17	90	6.4	432	D75110	hypothetical prote
18	87.5	6.2	508	T22836	surface protein -
19	87.5	6.2	1679	T30271	FKBP6 binding pro
20	87	6.2	802	A1227	phenylalanyl-trna
21	86.5	6.2	312	T25852	hypothetical prote
22	86.5	6.2	496	A41264	glucose transport
23	86.5	6.2	636	JC4960	DNA topoisomerase
24	86.5	6.1	299	G71088	hypothetical protei
25	85	6.0	408	H90222	GTP binding carboxyl
26	83.5	5.9	775	T40647	ubiquitin carboxyl
27	83.5	5.9	844	T23656	hypothetical prote
28	83	5.9	361	A75525	conserved hypothet

Query Match	Score	Query Match	Length	ID	Description
3	FEELLSQVGLGRFOMLHVF	14.0%	Score 197.5; DB 2; Length 576;		
17	FDVLEQVGVNGYQYQYFF	19.6%; Pred. No. 1.6e-09;			
56	NNTGSGNETGILSDALRIS	Best Local Similarity 19.6%; 51; Mismatches 90; Indels 159; Gaps 12;			
67	---PGRKE---YLR---PL				
114	---EPCVDGWNQSYFSTIY				
109	DRISLPCONGMDYNSIYD				
169	---KFGRRSEFVITLVLC				
169	---KFGRRSEFVITLVLC				
229	GRISGLMTSLFFGAMALL				
181	TKNLDGKALKRVARFNGI				
289	VGWKWDARKOLKIKKMG				
236	PSMRKRICIVFL 248				
345	PNLARKLLIYTYI 357				

ALIGNMENTS

RESULT 1
T22509
hypothetical protein F52F12.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22509
R:Matthews, L.
submitted to the EMBL Data Library, December 1996
A:Reference number: Z19573
A:Accession: T22509
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-576 <MID>
A:Cross-references: EMBL:Z83228; PIDN:CA05732.1; GSPDB:GN00019; CESP:F52F12.1
A:Experimental source: clone F52F12
A:Gene: CESP:F52F12.1
A:Map position: 1
A:Introns: 16/1; 23/3; 50/1; 80/3; 108/2; 134/3; 221/1; 257/2; 319/3; 358/2; 490/3; 5

conserved hypothet
high-affinity gluc
glucose transport
acrosomal protein
hypothetical prote
hypothetical prote
cholesterol monoox
hypothetical prote
probable voltage-a
probable cdh prote
proline-trna synth
hypothetical monoox
cholesterol monoox
protein C54E4.3 (1
MAPK-activated pro
glucose transport


```

OY 163 GGHVSDR----- 169
DB 167 VGYIADRGKRLCLLVTLVTVSGVLTAADYTMFLRLGMSKSGVSGYTLIT 226
OY 170 -----MLV-ES 174
DB 227 EFVSGYRRTAIIYQMAFYGVGLAGVAAVAPDMRMQLAVSLPTFLFLLYMFPES 286
OY 175 ARMLITNKLEDEKALKRKVARTNG-INKAEETLIEVVRSTMOEBLDAQTKT-TVCDL 232
DB 287 PFWLLSOKRTTRAVRIMQIAOKNGKVPAD-----LKMCLDEE-DASEKRSBSFADL 338
OY 233 FRNPSMRKRICILVF 248
DB 339 FRPTLRKRHTVILMYL 354

RESULT 5
JC4884
Organic cation transporter protein 2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1996 #sequence_revision 18-Oct-1996 #text_change 05-Nov-1999
C:Accession: J04884
R:Okuda, M.; Saito, H.; Urakami, Y.; Takano, M.; Inui, K.
Biochem. Biophys. Res. Commun. 224, 500-507, 1996
A:Title: cDNA cloning and functional expression of a novel rat kidney organic tra
A:Reference number: J04884; MUID:96295517
A:Accession: J04884
A:Molecule type: mRNA
A:Residues: 1-593 <OK>
A:Cross-references: DDBJ:D83044; NID:g1502282; PIDN:BAAL1754.1; PID:d1012421; PID:g15022
A:Experimental source: kidney
C:Comment: This protein is responsible for the transport of cationic drugs in kidney.

Query Match 11.6%; Score 162.5; DB 2; Length 593;
Best Local Similarity 18.3%; Pred. No. 2.1e-06;
Matches 68; Conservative 52; Mismatches 101; Indels 151; Gaps 12;

OY 4 EELLVVGGLGRQMLHLVFLPSLM-LIPIHLENFAALIPGHRCW----- 50
DB 5 DDLEHIGEPHFQ-KQFFELALLSGAFTPIYGVIGFTPDHHCSPGAATLSQRC 62
OY 51 -----VHMLDNNTSGNETGILSEDALEIRISIPLSNLRPKCRFRVHPOMQLHLNG 103
DB 63 GMSQAEELNVTYVGLSPDSASFSLQ--CMRYEV--DMNSTLDC--VDP-----LSS 109
OY 104 THISTSEADTEPCVDGWYDOSYFPSTIYTKMPLVCDYOSLKSVOFLLLTGMVGGIIG 163
DB 110 LAADRQQLPLGPEHGMYNTP--GSSIVTEFNLCAHSWMLDLFOSVAVVGFPGAMMI 167
OY 164 GHVSDR-----MLV----- 172
DB 168 GYLADRGKRLCLLVTLVTVSGVLTAADYTMFLRLGMSKSGVSGYTLIT 227
OY 173 -----ESA 175
DB 228 FVGLGYRRMVGICYQIAFTVGLLILAGVAVVIPNMRWLOFAVTLPNCFLLYFWCIPESP 287
OY 176 RMLITNKLEDEKALKRKVARTNG-INKAEETLIEVVRSTMOEBLDAQTKT-TVCDL 235
DB 288 RMLISOKRTTRAVRIMQIAOKNGKVPAD-----LKMCLDEE-DASEKRSBSFADL 338
OY 236 PSMRKRICILVF 247
DB 342 PDIRKHTVILMY 353

RESULT 6
T27870
Hypothetical protein ZK455.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

```

```

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T27870
R:White, S.
Submitted to the EMBL Data Library, November 1995
A:Reference number: Z20432
A:Accession: T27870
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-794 <MIL>
A:Cross-references: EMBL:Z66567; PIDN:CAA91492.2; GSPDB:GN00028; CESP:ZK455.8
A:Experimental source: clone ZK455
A:Genetics:
A:Gene: CESP:ZK455.8
A:Map position: X
A:Introns: 52/3; 119/2; 158/1; 207/3; 291/3; 335/2; 428/3; 610/2; 636/3; 733/3

Query Match 11.1%; Score 156; DB 2; Length 794;
Best Local Similarity 20.4%; Pred. No. 1.1e-05;
Matches 77; Conservative 46; Mismatches 103; Indels 152; Gaps 15;

OY 3 FEE-----LTSQVGLGRFQMLHLVFLPSLM-LIPIHLENFAAL-----PGRHCWVHML 54
DB 156 FEEDDGLIQLIGGCSYMQI--IVYLISVQ--QVPHAMF-NLSVYVMYQPDHWCRIPEF 211
OY 55 DNNTSGNETGILSE--DALLR--ISIPLDSN-----LRPKC-----RRVH-----PQWQ 97
DB 212 NERSFSA-ELGYNTYMDVOLNSTIAFPRTFNKORNELHNDQCHYERDVAHKLISFMAQ 270
OY 98 LHLN--GTIHSSEADTEPCVDGWYDOSYFPSTIYTKMPLVCDYOSLKSVOFLLLNG 155
DB 271 VKRMNATGKINRKE-----WEIDTSVMDRTIYTEMNRVODNMSAHVHMSYSLG 321
OY 156 MLVGGIIGHVSDRW----- 170
DB 322 YLVGCFVYGFISDRYGRKATIGTGLITMLFGFLTYSKFEIFLVYRFLAATNEAAL 381
OY 171 ----- 170
DB 382 AAYVLCMEVTGYRSTIVSLIQAPWAGYAFALALVILKRSWTMHLICVLAHIIISLM 441
OY 171 ---LVESARWLITNKLEDEKALKRKVARTNGIKNAEETLIEVVRSTMOEBLDAQTKT 227
DB 442 LYFLPESPRLIILNKKTKQAEKIRIACHYV--KSLRPSDLGLVRAEKKKWKHREKP 498
OY 228 TVCDLFRNPSMRKRICIL 245
DB 499 SYFHLFRSSELRFRNVYL 516

RESULT 7
T30895
sugar transport protein homolog T08B1.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 29-Oct-1999
R:Henkhaus, J.; Wohlmann, P.; Leimbach, D.
Submitted to the EMBL Data Library, August 1999
A:Description: The sequence of C. elegans cosmid T08B1.
A:Reference number: Z20926
A:Accession: T30895
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-591 <HEN>
A:Cross-references: EMBL:AF039039; PIDN:AAB94177.2
A:Experimental source: strain Bristol N2
C:Genetics:
A:Map position: V
A:Introns: 12/2; 62/1; 89/1; 111/3; 145/3; 325/2; 359/1; 468/2; 540/2
A:Note: T08B1.1

Query Match 7.7%; Score 108; DB 2; Length 591;

```


RESULT
F69587

Cross-references: EMBL:XL15864; NID:510190; PDB:5G1091
Hogan, A.; Heyner, S.; Charron, M.-J.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Development 113, 363-372, 1991

A:Title: Glucose transporter gene expression in early mouse embryos.
A:Reference number: A44887; MUID:92111400
A:Accession: B44887
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 384-431; P: 433-496 <HOG>
A:Cross-references: GB:S7926; NID:9242129; PIDN:AAB20847.1; PID:9242130
A:Experimental source: embryo
A>Note: sequence extracted from NCBI backbone (NCBI:77926, NCBI:77927)
R:Waelder, G.; Thompson, N.; Haefliger, J.A.; Nicod, P.
J. Biol. Chem. 269, 26912-26919, 1994
A:Title: Characterization of the murine high Km glucose transporter GLUT2 gene and its t
A:Reference number: A55078; MUID:95014557
A:Accession: I48367
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5 <RES>
A:Cross-references: EMBL:X78722; NID:9517486; PIDN:CAA55368.1; PID:9683589
C:Genetics:
A:Gene: GLUT2
C:Superfamily: glucose transport protein
C:Keywords: liver; transmembrane protein

Query Match 6.8%; Score 95; DB 2; Length 523;
Best local similarity 25.8%; Pred. No. 1.7;
Matches 34; Conservative 27; Mismatches 36; Indels 36; Gaps 6;

QY 143 SIKSVQFLITGMLVGGIIG-----GHVSDRWLV-----ESARW 177
DB 186 ALGTIHLALVTGILISQIAGLSFILGN-QDWHMLLGLSAVPALLQCLLLFCESPRY 244
QY 178 LIITNKLEDEGIALRKVARTNGINAEETLIEVVRSTMOEELDAQF--KTYCDLFRN 235
DB 245 LYL--NLEEEVRAKSKLRIGTETDITKDIN-----EMRKEKEASTEQKVSIVQLFTD 296
QY 236 PSMKRRIQIVFL 248
DB 297 ANYPQPIVVALML 309

RESULT 14
G84864
probable membrane transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: G84864
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
Nuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: G84864
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-521 <STO>
A:Cross-references: GB:AE002093; NID:92289003; PIDN:AAB64332.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g43330
A:Map position: 2
C:Superfamily: glucose transport protein

Query Match 6.8%; Score 93; DB 2; Length 521;
Best local similarity 26.9%; Pred. No. 2.6;
Matches 35; Conservative 23; Mismatches 38; Indels 34; Gaps 7;

QY 125 SYFSTIYTK---MDLVCDYOSLKSVPQFLITGMLVGGIIGHVSDRMLVESARWLI 180
DB 188 SYLVNSAFYQVGTWRMMLGVSGVPAVIOFILML-----FMPESPRLFLM 232
QY 181 TTKLDEGLAKLRVARTNGIKNAEETLIEVVRSTMOEELDAQOTKTTV--CDLFRNPSM 238

DB 233 KNKRAEAIQVL---ARTYDISRLD--EIDHLSAAEEEE---KORRTVGYLDVFRSKEL 284
QY 239 RKRICILVFL 248
DB 285 R-----LAFL 289

RESULT 15
A31556
glucose transport protein, hepatic - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 17-Nov-2000
C:Accession: A31556; S68362
R:Thorens, B.; Sarkar, H.K.; Kaback, H.R.; Lodish, H.F.
Cell 55, 281-290, 1988
A:Title: Cloning and functional expression in bacteria of a novel glucose transporter
A:Reference number: A31556; MUID:89003066
A:Accession: A31556
A:Molecule type: mRNA
A:Residues: 1-522 <THO>
A:Cross-references: GB:J03145; NID:9204534; PIDN:AAA41298.1; PID:9204535
R:Ahm, Y.; Kim, J.; Han, G.; Lee, B.; Kim, Y.
Arch. Biochem. Biophys. 323, 387-396, 1995
A:Title: Cloning and characterization of rat pancreatic beta-cell/liver type glucose
A:Reference number: S68362; MUID:96063615
A:Accession: S68362
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-246; K: 248-522 <AHN>
A:Cross-references: EMBL:L28134
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, September 199
C:Genetics:
A:Intons: 5/3; 36/3; 122/2; 164/1; 202/3; 257/1; 319/3; 354/3; 388/3; 456/3
C:Superfamily: glucose transport protein
C:Keywords: liver; transmembrane protein

Query Match 6.8%; Score 93; DB 2; Length 522;
Best local similarity 24.8%; Pred. No. 2.6;
Matches 33; Conservative 28; Mismatches 36; Indels 36; Gaps 6;
QY 143 SIKSVQFLITGMLVGGIIG-----GHVSDRWLV-----ESARW 177
DB 185 ALGTIHLALVTGILISQIAGLSFILGN-QDWHMLLGLSAVPALLQCLLLFCESPRY 243
QY 178 LIITNKLEDEGIALRKVARTNGINAEETLIEVVRSTMOEELDAQF--KTYCDLFRN 235
DB 244 LYL--NLEEEVRAKSKLRIGTETDITKDIN-----EMRKEKEASTEQKVSIVQLFTD 295
QY 236 PSMKRRIQIVFL 248
DB 296 PNYRQPIVVALML 308

Search completed: July 1, 2002, 11:41:47
Job time: 248 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2002, 14:01:31 ; Search time 316.64 Seconds
(Without Alignments)
4359.523 Million cell updates/sec

Title: US-09-674-235-10

Perfect score: 804
Sequence: 1 atgagccttgagagagctctt.....acacaaagtgcacaaattt 804

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: N_Geneseq_032802:*
 - 1: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT.*
 - 2: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT.*
 - 3: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT.*
 - 4: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT.*
 - 5: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT.*
 - 6: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT.*
 - 7: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT.*
 - 8: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT.*
 - 9: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT.*
 - 10: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT.*
 - 11: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT.*
 - 12: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT.*
 - 13: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT.*
 - 14: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT.*
 - 15: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT.*
 - 16: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT.*
 - 17: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT.*
 - 18: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT.*
 - 19: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT.*
 - 20: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT.*
 - 21: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT.*
 - 22: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT.*
 - 23: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT.*
 - 24: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	804	100.0	804	21	AAZ38316 Human transmembran
2	804	100.0	1705	21	AAZ38317 Human transmembran
3	782	97.3	1030	22	AAH98847 Human EST-derived
4	508	63.2	1353	22	ABA08976 Human secreted pro
5	508	63.2	2684	22	AAH5824 HOAT5 DNA. Homo s
6	335.2	41.7	1375	22	AAI59215 Human polynucleoti
7	335.2	41.7	1986	22	AAAD09561 Human transporter
8	332	41.3	2027	21	AAAC1892 CDNA encoding a hu
9	314.2	39.1	1977	22	AAAC5823 HOAT4 DNA. Homo s

10	232.8	29.0	1132	22	AAI61001 Human polynucleoti
11	208.8	26.0	2455	22	AAD17476 Human transporter
12	208.8	26.0	2865	22	AAD12566 Human protein havi
13	180.8	22.5	2047	22	AAD17480 Human transporter
14	180.8	22.5	2194	22	AAD12607 Human protein havi
15	180.8	22.5	2210	22	AAH32614 Human organic anio
16	160.4	20.0	540	22	ABA63144 Human foetal liver
17	160.4	20.0	540	22	ABA30394 Probe #8860 for ge
18	160.4	20.0	540	22	AAK11579 Human brain expres
19	160.4	20.0	540	22	AAK37346 Human bone marrow
20	160.4	20.0	540	22	AAI18174 Probe #18107 for ge
21	160.4	20.0	540	22	AAI43191 Probe #11877 used
22	145.2	18.1	447	22	AAI07074 Human reproductive
23	145.2	18.1	9370	22	AAI07075 Human reproductive
24	136.2	16.9	339	22	ABA75594 Human foetal liver
25	136.2	16.9	339	22	ABA40206 Probe #18672 for g
26	136.2	16.9	339	22	AAK24178 Human brain expres
27	136.2	16.9	339	22	AAK50227 Human bone marrow
28	136.2	16.9	339	22	AAI27308 Probe #17241 for g
29	136.2	16.9	339	22	AAI56179 Probe #24865 used
30	134.4	16.7	685	22	AAH98991 Rat EST-derived co
31	130.8	16.3	1316	22	AAH55043 Nucleotide sequenc
32	108	13.4	789	22	AAK92269 Human CDNA 5'-end
33	108	13.4	789	22	AAK93534 Human CDNA clone r
34	108	13.4	1871	22	AAK94561 Human full-length
35	98.2	12.2	419	22	AAI02433 Human reproductive
36	88.8	11.0	280	21	AAAI0196 Rat liver toxicolo
37	88.2	11.0	2294	20	AAV79584 Rat organic anion
38	85.2	10.6	379	22	AAI84356 Human polynucleoti
39	81.8	10.2	2123	21	AAZ29300 Human organic anio
40	81.8	10.2	2127	22	AAH9585 HOAT1 DNA. Homo s
41	81.8	10.2	2171	20	AAV79585 Human organic anio
42	81	10.1	1638	22	AAH80044 Mouse organic anio
43	79.4	9.9	596	22	AAI82398 Human polynucleoti
44	74	9.2	563	22	AAK92176 Human CDNA 5'-end
45	74	9.2	563	22	AAK93441 Human CDNA clone r

ALIGNMENTS

RESULT 1	AAZ38316	standard; CDNA; 804 BP.
ID	AAZ38316	
XX	AAZ38316	
AC	AAZ38316	
XX	09-FEB-2000	(first entry)
DT	XX	
XX	XX	
DE	XX	Human transmembrane protein CDNA clone HP02000 coding sequence.
XX	XX	
KW	HP02000	: transmembrane domain; liver; expression; homology;
KW	organic cation transporter; drug excretion; antibody; assay reagent;	
KW	diagnostic marker; primer; probe; antisense; gene therapy;	
KW	agonist; antagonist; ligand; therapeutic; ds.	
XX	XX	
OS	XX	Homo sapiens.
XX	XX	
FH	XX	
FT	Key	Location/Qualifiers
FT	CDS	1..804
FT	/*tag=	a
FT	/product=	"Human transmembrane protein HP02000"
FT	/note=	"No stop codon given in the specification"
XX	XX	
XX	MO9955862-AZ.	
PN	XX	
XX	XX	
PD	04-NOV-1999.	
XX	XX	
XX	27-APR-1999;	99MO-JP02226.
PF	XX	
XX	28-APR-1998;	98UP-0119395.
PR	XX	
XX	XX	
PA	(SAGA)	SAGAMI CHEM RES CENT.

PA (PROT-) PROTEGENE INC.

XX Kato S, Kimura T;

XX WPI: 2000-023358/02.

DR P-PSDB; AAY52386.

XX Human proteins with transmembrane domains, involved in control of cell
 proliferation and differentiation, useful for treating e.g. cancer or
 inflammation

PS Claim 3; Page 84; 114pp; English.

XX This sequence represents the coding sequence of human CDNA clone
 CC HP02000 which encodes a 32 kD protein with two putative transmembrane
 CC domains. The CDNA was isolated from a human liver cell CDNA library, and
 CC from tissue localisation studies has been found to be expressed only in
 CC the liver. The protein has homology with the rat organic cation
 CC transporter (EMBL Accession No. Y09945) which is involved in drug
 CC excretion, and may have a similar function. The protein
 CC may be used to raise specific antibodies, as assay reagents, as
 CC diagnostic tissue markers, for the isolation of cognate receptors,
 CC ligands and binding proteins, and as biologically active agents.
 CC Nucleotides encoding the protein may be used as primers and probes or
 CC antisense molecules, and in gene therapy. Cells transformed with these
 CC nucleotides may be used to screen for agonists and antagonists which are
 CC potentially useful therapeutically.

XX Sequence 804 BP; 226 A; 171 C; 193 G; 214 T; 0 other;

Query Match 100.0%; Score 804; DB 21; Length 804;
 Best Local Similarity 100.0%; Pred. No. 1.9e-248;

Matches 804; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggccttggagagctcttgatgaagtgtggagccttgggaatttcagatctcat 60
 DB 1 atggccttggagagctcttgatgaagtgtggagccttgggaatttcagatctcat 60
 QY 61 ctggtttattcttcctccctccatgtttaaccctataactcgtgagaacttt 120
 DB 61 ctggtttattcttcctccctccatgtttaaccctataactcgtgagaacttt 120
 QY 121 gctgcagcattcctcgtgcatcgttgcgtgggtccacatgctgacaaataactgac 180
 DB 121 gctgcagcattcctcgtgcatcgttgcgtgggtccacatgctgacaaataactgac 180
 QY 181 ggttaagaactcgtgactcctcagtgagaatgacctcttgagaactctatccactagac 240
 DB 181 ggttaagaactcgtgactcctcagtgagaatgacctcttgagaactctatccactagac 240
 QY 241 tcaaatctgagcagcagagaagtgcgtcgttgcattccaccagtgagcactcttcac 300
 DB 241 tcaaatctgagcagcagagaagtgcgtcgttgcattccaccagtgagcactcttcac 300
 QY 301 ctgaatggagactaccacagcacaagtgaagcagacacagacccttgtgtagtctg 360
 DB 301 ctgaatggagactaccacagcacaagtgaagcagacacagacccttgtgtagtctg 360
 QY 361 gtaatatgacaaagctactccttcgacacatgtgactaagtggagcctgtgtgat 420
 DB 361 gtaatatgacaaagctactccttcgacacatgtgactaagtggagcctgtgtgat 420
 QY 421 tatcagtcactgaatcagtggttcaattcctactcttcgtgactgaaatgctggaggagc 480
 DB 421 tatcagtcactgaatcagtggttcaattcctactcttcgtgactgaaatgctggaggagc 480
 QY 481 atcatatggcgcacatgtctcaagcaggtgctgtggaactcgtcgtgttatatac 540
 DB 481 atcatatggcgcacatgtctcaagcaggtgctgtggaactcgtcgtgttatatac 540
 QY 541 accaataaactgagtgaagccttaagcacttagaaagtgtgacgcacaaatgagata 600
 DB 541 accaataaactgagtgaagccttaagcacttagaaagtgtgacgcacaaatgagata 600

DB 541 accaataaactgagtgaagccttaagcacttagaaagtgtgacgcacaaatgagata 600
 QY 601 aagaatgctgaagaagacccttgacatagaggtgtgtaagatccaccatgcaggagctg 660
 DB 601 aagaatgctgaagaagacccttgacatagaggtgtgtaagatccaccatgcaggagctg 660
 QY 661 gatgcagcacagacaaactactgtgtgactgttccgcacccagtatgcgttaa 720
 DB 661 gatgcagcacagacaaactactgtgtgactgttccgcacccagtatgcgttaa 720
 QY 721 aggatcgtatctcgtgatttttgagaaaaaatctcaaggaagagcataaatgat 780
 DB 721 aggatcgtatctcgtgatttttgagaaaaaatctcaaggaagagcataaatgat 780
 QY 781 tgcacacaaagtgcacaaattt 804
 DB 781 tgcacacaaagtgcacaaattt 804

RESULT 2

ID AA238317 standard; CDNA; 1705 BP.

XX AA238317;

DT 09-FEB-2000 (first entry)

XX Human transmembrane protein CDNA clone HP02000.

KW HP02000; transmembrane domain; liver; expression; homology;

KW organic cation transporter; drug excretion; antibody; assay reagent;

KW diagnostic marker; primer; probe; antisense; gene therapy;

KW agonist; antagonist; ligand; therapeutic; ds.

OS Homo sapiens.

FT Key

FT CDS

FT FT

FT FT

PN W09955862-A2.

PD 04-NOV-1999.

PF 27-APR-1999; 99WO-JP02226.

PR 28-APR-1998; 98JP-0119395.

PA (SAGA) SAGAMI CHEM RES CENT.

PA (PROT-) PROTEGENE INC.

PI Kato S, Kimura T;

XX WPI: 2000-023358/02.

DR P-PSDB; AAY52386.

XX Human proteins with transmembrane domains, involved in control of cell

XX proliferation and differentiation, useful for treating e.g. cancer or

XX inflammation

XX Claim 4; Page 89-91; 114pp; English.

XX This sequence represents the human CDNA clone HP02000 which
 CC encodes a 32 kD protein with two putative transmembrane
 CC domains. The CDNA was isolated from a human liver cell CDNA library, and
 CC from tissue localisation studies has been found to be expressed only in
 CC the liver. The protein has homology with the rat organic cation
 CC transporter (EMBL Accession No. Y09945) which is involved in drug
 CC excretion, and may have a similar function. The protein
 CC may be used to raise specific antibodies, as assay reagents, as
 CC diagnostic tissue markers, for the isolation of cognate receptors,
 CC ligands and binding proteins, and as biologically active agents.

QY 301 ctgaatggagctataccacagcaaaagtgagcagacacagaccctgtgtgtagtgcg 360
 |||||||
 Db 486 ctgaatggagctataccacagcaaaagtgagcagacacagaccctgtgtgtagtgcg 545
 |||||||
 QY 361 gatatgtatcaaaagctacttcccttcgaccattgtgtaagtgtggagcctgtatgtat 420
 |||||||
 Db 546 gtatgtatcaaaagctacttcccttcgaccattgtgtaagtgtggagcctgtatgtat 605
 |||||||
 QY 421 tatcagctagaatcaatggtgttcacatccctactctactctgactggaatgctgggtggagc 480
 |||||||
 Db 606 taccagctagaatcaatggtgttcacatccctactctactctgactggaatgctgggtggagc 665
 |||||||
 QY 481 atcataagtgccatgctcagacacaggtgctgtggaatctgctgctgtgtgataatc 540
 |||||||
 Db 666 atcataagtgccatgctcagacacaggtgctgtggaatctgctgctgtgtgataatc 725
 |||||||
 QY 541 accataaactagatgagggcttaagagcacttagaagaagtgtcagcagacaatgtgata 600
 |||||||
 Db 726 accataaactagatgagggcttaagagcacttagaagaagtgtcagcagacaatgtgata 785
 |||||||
 QY 601 aagaatgctg-aagaacccctgaacacatagaggtgtgaaatccacatcagcagagagct 659
 |||||||
 Db 786 aagaatgctgaaagaacccctgaacacatagaggtgtgaaatccacatcagcagagagct 845
 |||||||
 QY 660 ggaatgagcagacagacacaaactactgtgtgtgactgtgtccgacacccagatgctgtaa 719
 |||||||
 Db 846 ggaatgagcagacagacacaaactactgtgtgtgactgtgtccgacacccagatgctgtaa 905
 |||||||
 QY 720 aagatcgtatcctcgtgattttttag-aaaaaaatcctaaggaaagcataaaatg 778
 |||||||
 Db 906 aagatcgtatcctcgtgattttttagaaaaaaatcctaaggaaagcataaaatg 965
 |||||||
 QY 779 atgtctacacaaaagtgaccaaattt 804
 |||||||
 Db 966 atgtctacacaaaagtgaccaaattt 991
 |||||||

RESULT 4

ABA08976 standard; cDNA; 1353 BP.

XX ABA08976;

DT 11-JAN-2002 (first entry)

DE Human secreted protein homologue-encoding cDNA, SEQ ID NO:752.

XX Human: cytokine; cell proliferation; cell differentiation; growth factor;
 KW hematoopoiesis regulation; tissue growth; immunomodulatory; activin;
 KW inhibitor; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferation; retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiautistic; haemostatic; antiarteriosclerotic;
 KW cytosolic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antilucer; ss.

OS Homo sapiens.

XX MO200157188-A2.

PD 09-AUG-2001.

XX 05-FEB-2001; 2001MO-US03800.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

XX (HYSF-) HYSFO INC.

XX Tang YT, Liu C, Dzumanac RT;
 XX WPI; 2001-457740/49.
 DR P-PSDB; ABB11732.
 DR
 XX
 PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 XX
 PS Claim 1; Page 677; 1963pp; English.

XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness.
 CC Autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention.

SQ Sequence 1353 BP; 349 A; 298 C; 315 G; 391 T; 0 other;

Query Match 63.28; Score 508; DB 22; Length 1353;
 Best Local Similarity 100.0%; Pred. No. 6; 2e-153;
 Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggcctttagagagccttgatcgaagttgagagccttgggaatttcagatctcat 60
 |||||||
 Db 186 atggcctttagagagccttgatcgaagttgagagccttgggaatttcagatctcat 245
 |||||||
 QY 61 ctggttttatcttccctctcctcatgtatlaatccctcatatcgtcgaagaactt 120
 |||||||
 Db 246 ctggttttatcttccctctcctcatgtatlaatccctcatatcgtcgaagaactt 305
 |||||||
 QY 121 gctcagcaccatcctcgtgcatcgttctggtgtccacatgctggaacaataactgagctc 180
 |||||||
 Db 306 gctcagcaccatcctcgtgcatcgttctggtgtccacatgctggaacaataactgagctc 365
 |||||||
 QY 181 ggaatgaacttgaaatccctcagtgaaatgagcctcttggaactctatccactgac 240
 |||||||
 Db 366 ggaatgaacttgaaatccctcagtgaaatgagcctcttggaactctatccactgac 425
 |||||||

XX 2

of diseases associated with

680 atcataggtggccatgttctcagacaggt 707

PR 19-JUL-2000; 2000US-0653450.
PP 03-AUG-2000; 2000US-0653450.

PR 19-JUL-2000; 2000US-0653450.
PR 03-AUG-2000; 2000US-0653450.

xx (HISE-) HYSEQ INC.
xx
xx
xx Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB,
PI Wang J, Wang Z, Wehman T, Xu C, Xie A, Ren F, Wang D,
PI Zhao Q, Zhou P, Goodrich R, Dimaec RT, Zhang J;
xx
xx WPI, 2001-44223/47.
xx P-PSDB, AAm40059.
xx

over, nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries -
xxx
xxx
xxx
Claim 1; SEQ ID NO 1418; 10078bp; English.
C
mbo-4

the invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AA836642-AA442213) with neurotropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Cytotoxic/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders.

Note: The sequence data for this patent did not form part of the printed specification.

Sequence 1375 BP; 354 A; 329 C; 324 G; 368 T; 0 other;

Query Match	41.7%;	Score 335.2;	DB 22;	Length 1375;
Best Local Similarity	78.7%;	Pred. No. 3e-97;		
Matches 400;	Conservative	0;	Mismatches	100

	Accession	Index	Gaps
QY	1 atgaccttggaggagcctctgtatgcgaagtctggagacctggagacttgagatccatgcagctcat	60	
Db	288 atgaccttggaggagcctccctgggacacagctgggagactctgagagatccatgcagctcat	60	
QY	61 ctggtttttttctccctccctccatgltatctatcccatctactgtatgaagactt	120	
Db	348 actgtttttcccatcactctgtctgtgtacatactcatctcatcttattatgttggaagactt	120	
QY	121 gctgcagacatctctgtgcatactgtctctgtggccacatgtctgacacataatactgacact	407	
Db	408 actgcattcatactatcgtgcacatcgtctgtgcccacatccctggacaaatgacacactctct	180	
QY	181 gtaatacgaactcgtgaatccctcagtggaagatgcctcttgagaatctcatccactagac	240	
Db	468 gacaaatgacacctggggcctccctccacgaagatgcacactcttgagaatccatccatccactgac	527	
QY	241 tccaactctaggccagcagaagaatgctgcctgttgcatacccccagtgagagctttctcac	300	
Db	528 tccaacatgagccagcagaagaatgctgcctgttgcatacccccagtgagagctttctcac	300	
QY	301 ctggaatgggactatccacagcacaaagtgaagccacacacagaacctcgtgtgagatgctg	360	
Db	568 ctggaatgggactatccacacacaaagtgaagccacacacagaacctcgtgtgagatgctg	360	
QY	361 gtaatacgaacactcactccctccctgacacatgtggaactaagtgaggacactgttatgtat	420	
Db	648 gtaatacgaacactcactccctccctcattccacacatgtgacgtgagatcgtgtatgtac	707	
QY	421 tatcagtcacatgaatataagtggttccaactctactcactctacacggagatcgtgtggaagc	480	
Db	708 tccaacactcagctccagtcgttaaatattgtattcatcatgctcggagatataatgtgtggaagc	767	

Qy 481 atcatagtgtygcattgtctcagacagt 508
||| |||| | ||| | ||||| |||
Db 768 atcctagcggtcatttattcagacagt 795

RESULT 7
AAD09561

AC AAD09561;
XX

10 SEP 2001 (first entry)
XX

Human transporter and ion channel-10 (TRICH-10) cDNA.

gene therapy; amphotrophic lateral sclerosis; amnesia; muscular dystrophy; hypertension; angina; neurological disorder; asthma; bipolar disorder; dementia; depression; Alzheimer's disease; epilepsy; vaccine; arthritis; Pick's disease; ischemic cerebrovascular disease; AIDS; anxiety; stroke; Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy; demyelinating disease; mental disorder; Schizophrenia; lupus; myositis; muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease; dermatomyositis; diabetes mellitus; immunological disorder; psoriasis; rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus; sickle cell anaemia; Wilson's disease; infertility; Cushing's disease; scleroderma; pulmonary artery stenosis; nootropic; Addison's disease; malabsorption syndrome; hypercholesterolemia; Cancer; ss.

T		location/Qualifiers	
T	CDS	248..1909	
F		/tag= a	
K		/product= "Human TRICH-10 protein"	
N	MOZ00146258-A2.		

28-JUN-2001.
22-DEC-2000; 2000WO-US35095.

23-DEC-1999; 9905-0172000
14-JAN-2000; 200005-0176083.
21-JAN-2000; 200005-0177332.
28-JAN-2000; 200005-0178572.
02-FEB-2000; 200005-0179758.
10-FEB-2000; 200005-0181625.

(INCY-) INCYTE GENOMICS INC.

Baughn MR, Burford N, Au-Young J, Lu DAM, Yang J, Reddy R, Lal P, Hillman JL, Azimzai Y, Yue H, Nguyen DB, Yao MG, Ganchi AR, Tang YT, Khan FA;

WPI; 2001-418042/44.
P-PSDB; AAE04897.

Novel human transporter and ion channel proteins useful for treating and preventing transport, neurological, muscle and immunological disorders -

Claim 5; Page 147; 160pp; English.

CC The present sequence is transporter and ion channel-10 (TRICH-10) cDNA
CC TRICH is used as vaccine. TRICH is useful for treating a disease or
CC condition associated with decreased expression of functional TRICH,
CC such as transporter disorder including amphotropic lateral sclerosis,
CC cystic fibrosis, Becker's muscular dystrophy Charcot-Marie Tooth
CC disease, Duchenne muscular dystrophy, angina and hypertension,
CC neurological disorders including Alzheimer's disease, and hypertension,
CC disorder, dementia, depression, epilepsy, ischaemic cerebrovascular

CC disease, stroke, cerebral neoplasms, Pick's disease, Huntington's
 CC disease and Parkinson's disease, demyelinating diseases, mental disorders
 CC including mood, anxiety, schizophrenia and seasonal affective disorder,
 CC muscle disorder including cardiomyopathy, myocarditis, polymyositis,
 CC dermatomyositis, arrhythmias and asthma and immunological disorders
 CC including AIDS, adult respiratory distress syndrome (ARDS), allergies,
 CC anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's
 CC syndrome, systemic lupus erythematosus and other diseases including
 CC sickle cell anaemia, Wilson's disease, cataracts, infertility, pulmonary
 CC artery stenosis, Grave's disease, Cushing's disease, Addison's disease,
 CC glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers
 CC psoriasis and viral, bacterial, fungal, helminthic and protozoal
 CC infections. TRICH DNA is useful in gene therapy and in diagnostic
 CC purposes.

Sequence 1986 BP; 505 A; 496 C; 457 G; 528 T; 0 other;

Query Match 41.7%; Score 335.2; DB 22; Length 1986;
 Best Local Similarity 78.7%; Pred. No. 3.7e-97;
 Matches 400; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 1 atggccttggagagcctcttgatgaagcttggagccttggagagattcaagcttcat 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 248 atggccttggagagcctcttgatgaagcttggagccttggagagattcaagcttcat 307
 QY 61 ctggtttttttcttcctctcctcatgtatatacctcctatatactctgtagaacttt 120
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 308 actgttttcttcaatctctgtctgtgtacatactctcttattatgtggaacttc 367
 QY 121 gctcagcattccctgtgcatcgtctgtgtgtccacatgctgagacaataactgact 180
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 368 actgcatctatactcctgctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 427
 QY 181 ggtatgaagcttgaatcctcagtggaatgcccctcttgagaactctatccactagac 240
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 428 gaaatgacctggggcctcagcagcaagatgacctctggaatctccatccactgagac 487
 QY 241 tcaaatctggagcagagagatgctgcgtctgtgtgtgtgtgtgtgtgtgtgtgtgt 300
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 488 tcaaatctggagcagagagatgctgcgtctgtgtgtgtgtgtgtgtgtgtgtgtgt 547
 QY 301 ctgaatggagcattccacagcagagtgagagacagacacacctgtgtgtagtgctg 360
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 548 ctgaatggagcattccacagcagagtgagagacagacacacctgtgtgtagtgctg 607
 QY 361 gttatgataaagctactcctcctgacacattgtactaaagtggagacctgtagtgc 420
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 608 gttatgataaagctactcctcctgacacattgtactaaagtggagacctgtagtgc 667
 QY 421 tatcagtcactgaatgaatggtgtcattctcctactctgactggaatctggtggaagc 480
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 668 tctcaactactgacttgaatggtggaatctcctactctgactggaatctggtggaagc 727
 QY 481 atcataggttgccatgctcagacagct 508
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 728 atctcaggcggtcattatcatgacagct 755

RESULT 8
 AAC61892
 ID AAC61892 standard; cDNA; 2027 BP.

AC AAC61892;

DT 06-MAR-2001 (first entry)

DE cDNA encoding a human secreted protein.

KX Secreted protein; platelet disorder; stem cell disorder; osteoporosis;
 KW osteoarthritis; burn; incision; ulcer; periodontal disease; neuropathy;
 KW nervous system disease; bone growth; cosmetic plastic surgery;
 KW gut protection; gut regeneration; fibrosis; cancer;

KW bone marrow transplantation; ss.

OS Homo sapiens.

FT Key Location/Qualifiers

FT CDS 270..1289

FT /tag= a /product= "secreted protein"

WT MO200061755-A2.

PD 19-OCT-2000.

PF 10-APR-2000; 2000MO-0509555.

PR 09-APR-1999; 99US-0128574.

PR 20-AUG-1999; 99US-0150054.

XX (CHIR) CHIRON CORP.

XX Garcia PD;

XX WPI: 2000-665133/64.

XX P-PSDB; AAB19406.

XX Novel secreted human proteins useful for stimulating blood cell

XX generation in patients receiving cancer chemotherapy, treating bone

XX marrow transplantation patients and for healing fractured bones

XX Claim 2; Page 71-72; 74pp: English.

XX AAC61879-93 encode secreted human proteins. The secreted proteins are

XX useful in assays to determine their biological activities. The proteins

XX can also be used as biomarkers to identify tissues or cell types which

XX express the proteins. The polynucleotide molecules can be used as

XX biomarkers for tissues or chromosomes and to elicit immune responses.

XX The proteins and antibodies are useful in diagnosis and treatment of

XX diseases associated with altered expression of these proteins. The

XX proteins are also useful for prevention or treatment of platelet

XX disorders, stem cell disorders, osteoporosis or osteoarthritis, burns,

XX incisions, ulcers, periodontal diseases, central and peripheral nervous

XX system diseases and neuropathies, for healing fractured bones and to

XX induce cartilage and/or bone growth in cosmetic plastic surgery. The

XX proteins are also useful for gut protection or regeneration, for the

XX treatment of lung or liver fibrosis, for stimulating blood cell

XX generation in patients receiving cancer chemotherapy and for treatment

XX of bone marrow transplantation patients.

XX Sequence 2027 BP; 530 A; 501 C; 459 G; 537 T; 0 other;

Query Match 41.3%; Score 332; DB 21; Length 2027;
 Best Local Similarity 78.3%; Pred. No. 4.1e-96;
 Matches 398; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 1 atggccttggagagcctcttgatgaagcttggagccttggagagattcaagcttcat 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 270 atggccttggagagcctcttgatgaagcttggagccttggagagattcaagcttcat 329
 QY 61 ctggtttttttcttcctctcctcatgtatatacctcctatatactctgtagaacttt 120
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 330 actgttttcttcaatctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 389
 QY 121 gctcagcattccctgtgcatcgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 180
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 390 actgcatctatactcctgcatcgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 449
 QY 181 ggtatgaagcttgaatcctcagtggaatgagcctcttgagaactctatccactagac 240
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 450 gcaatgagacatgggcccacagcagatgacatcttgagaactctcaccaccactgac 509
 QY 241 tcaaatctaggcagagagatgctgcgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

PR 29-NOV-2000; 2000US-0727344.
 XX
 XX (HXSE-) HXSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 XX Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 XX Zhao QH, Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 DR P-PSDB; AAM41845.
 XX
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PR such as central nervous system injuries -
 XX
 XX Claim 1: SEQ ID NO 4990; 10078BP; English.
 PS
 XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with neurotropic,
 CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombotic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 XX Sequence 1132 BP; 274 A; 265 C; 285 G; 305 T; 3 other;
 SO
 Query Match 29.0%; Score 232.8; DB 22; Length 1132;
 Best Local Similarity 83.5%; Pred. No. 2.7e-64;
 Matches 264; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
 QY 193 ggaatcctcagtgagatgagcctcttggaatcctctccactagactcaactgag 252
 DB 27 ggggcccacagcaagatgacactcttggaatcctccactgagactcaactgag 86
 QY 253 ccagagaagtgctgctgcttgcattccacagtgagcctctcactgaatggag 312
 DB 87 ccagagaagtgctgctgcttgcattccacagtgagcctctcactgaatggag 146
 QY 313 atccacagacagatgagagcagacacagacccctgtgtgagtgctggtatagataa 372
 DB 147 tcccccacacagatgagcagcagacatgagccctgtgtgagtgctggtatagataa 206
 QY 373 agctactcctcctgacacattgagtgagtgagcctgtagatgattatcaagcactg 432
 DB 207 atctccttctcaccacacctcgtgagtgagtgagcctgtagatgattatcaagcactg 266
 QY 433 aaatgagtgctcaacttacttactgagtgagtgagcctgtagatgattatcaagcactg 492
 DB 267 acttcagtgagcctaattatctatctatgctgtagatgattatgagcactgagcgt 326
 QY 493 catgtctcagacaggt 508
 DB 327 cattatcagacaggt 342
 RESULT 11
 AAD17476 standard: cDNA; 2455 BP.
 XX AAD17476;
 AC AAD17476;
 XX
 DT 10-DEC-2001 (first entry)
 XX

DE Human transporter and ion channel-9 (TRICH-9) cDNA.
 XX Human; transporter and ion channel; TRICH-9; therapy; akinesia; cardiast;
 KW neurological disorder; immune disorder; allergy; neurotropic; dementia;
 KW AIDS; Acquired Immune Deficiency Syndrome; amyotrophic lateral sclerosis;
 KW cystic fibrosis; Alzheimer's disease; depression; epilepsy; scleroderma;
 KW cardiomyopathy; muscular dystrophy; Crohn's disease; multiple sclerosis;
 KW rheumatoid arthritis; cancer; anaemia; asthma; trauma; haemodialysis;
 KW gastritis; inflammation; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH 2..1663
 FT CDS /tag= a
 FT /product= "Human TRICH-9 protein"
 XX
 XX WO200162923-A2.
 PN
 XX 30-AUG-2001.
 XX
 XX 23-FEB-2001; 2001WO-US05942.
 XX
 XX 25-FEB-2000; 2000US-0184866.
 PR 02-MAR-2000; 2000US-0187947.
 PR 09-MAR-2000; 2000US-0188333.
 PR 17-MAR-2000; 2000US-0190230.
 PR 24-MAR-2000; 2000US-0192077.
 PR 30-MAR-2000; 2000US-0193500.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Yue H, Tang YT, Lal P, Policky JL, Nguyen DB, Au-Young J, Yao MG;
 PI Khan FA, Walla NK, Gandhi AR, Tribouley CM, Patterson C;
 PI Thornton M, Greene BD, Hernandez R, Borowsky ML, Sanjanna MS;
 XX
 XX WPI: 2001-582050/65.
 DR P-PSDB; AAE10332.
 XX
 XX Thirteen human transporters and ion channels (referred to as TRICH-1 to
 PT TRICH-13), useful in the diagnosis, treatment and prevention of
 PT transport (e.g. akinesia), neurological, muscle or immunological
 PT disorders (e.g. allergies) -
 XX
 XX Claim 11: Page 127-128; 131pp; English.
 PS
 XX The present sequence is a cDNA encoding human transporters and ion
 CC channels (TRICH-9) protein. The TRICH DNA, protein and their agonist and
 CC antagonists are useful in the diagnosis, treatment and prevention of
 CC transport disorders (akinesia, amyotrophic lateral sclerosis, cystic
 CC fibrosis), neurological (Alzheimer's disease, dementia, depression,
 CC epilepsy), muscle (e.g. cardiomyopathy, Becker's muscular dystrophy) or
 CC immunological disorders (e.g. allergies, acquired immunodeficiency
 CC syndrome (AIDS), Crohn's disease, acquired immunodeficiency
 CC multiple sclerosis), viral, bacterial, parasitic, scleroderma,
 CC hematologic disorders, cancer, anaemia, asthma, trauma, haemodialysis,
 CC gastritis and inflammation.
 CC
 XX Sequence 2455 BP; 416 A; 842 C; 707 G; 490 T; 0 other;
 SO
 Query Match 26.0%; Score 208.8; DB 22; Length 2455;
 Best Local Similarity 63.2%; Pred. No. 2.3e-56;
 Matches 321; Conservative 0; Mismatches 187; Indels 0; Gaps 0;
 QY 1 atggccttgagagagcttgagcagtgagagccttgagagattcagatgctcat 60
 DB 2 atggcattctcgaactcctgagcctggtggtgagcctgagagattcagatgctcat 61
 QY 61 ctgatttattctcctctctcattatataacccctcattatcgtcagagactt 120
 DB 62 acgattgctctgattgctcattcattggtgctgttaccacagagatgctgagagactt 121

QY 121 gctgcagccattctcgtgctcgtgtggtggtccacatgctgagcaataactgattct 180
 Db 122 tcgcccgcgcgtccacagccagcgtgctggtgacccctccgagcaacacagcagctcag 181
 QY 181 ggtatgaacttgatcccaatgagatgacctcttgaaatctctatcccaagtagac 240
 Db 182 gccagagatcctagggagctgagctcctcctgagccctcctgatttccatccgcgggc 241
 QY 241 tcaatctgagggcagagagatgctgctggtggtggtccacagggagcctctacac 300
 Db 242 cccaaccagagagcccccagcagtcgctgctccgcccagcagagtgagcagctctgag 301
 QY 301 ctgaatggagactatccacagcagcaagtgagggagacacagaccctggtgagtgctg 360
 Db 302 cccaatgagcagggccacagctgagagcgagcgagcagagagcggtgtgtgagtgctg 361
 QY 361 gatatgatcaagaactcctcctcctgagcattgagcaagtggagcagctggtatgtat 420
 Db 362 gctatgacgcagcattctcctccacacatcggtgagcagatggagacctcgtgtgagc 421
 QY 421 tatcagtaactgaatcaatggttcaatcctcctcctgagcagtgagtggtgagagc 480
 Db 422 tctcagctctgaaagccatgagccagctcctcctcctcctcctcctcctcctcctcct 481
 QY 481 atcataagtgccatgctcctcctcctcctcctcctcctcctcctcctcctcctcctc 508
 Db 482 gctgctgagggccctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 509

RESULT 12

AAD12566
 ID AAD12566 standard; CDNA; 2865 BP.

AC AAD12566;

DT 25-SEP-2001 (first entry)

XX Human protein having hydrophobic domain encoding cDNA clone HP03613.

KW Human; hydrophobic domain; gene therapy; nutritional supplement;
 KW cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;
 KW multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;
 KW haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;
 KW Huntington's disease; Alzheimer's disease; chemotactic; thrombotic;
 KW haemostatic; thrombolytic; tumour growth inhibitor; anabolic;
 KW contractile; antinflammatory; antinflammatory; ss.

XX Homo sapiens.

OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 338..2074
 FT /tag= a
 FT /product= "Human protein having hydrophobic domain"
 FT /note= "CDS is specifically is claimed in claim 3"

PN WO200149728-A2.

PD 12-JUL-2001.

PE 28-DEC-2000; 2000WO-JP03359.

PR 06-JAN-2000; 2000JP-0000585.

PR 11-JAN-2000; 2000JP-0000588.

PR 03-FEB-2000; 2000JP-0002299.

PR 03-MAR-2000; 2000JP-0058367.

PA (PROT-) PROTEGENE INC.

PA (SAGA) SAGAMI CHEM RES CENT.

PI Kato S, Kimura T;

XX WPI; 2001-41835/44.

DR P-PSDB; AAE06571.
 XX
 PT Human proteins with hydrophobic domains and the nucleic acids encoding
 PT them, useful for preventing diagnosing and treating e.g. cancer,
 PT Alzheimer's and inflammation -
 PS Claim 4; Page 252-257; 563pp; English.

The present sequence is human protein with hydrophobic domain encoding
 CC cDNA clone HP03613. The polynucleotide and polypeptide of the invention
 CC may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate polypeptide expression. The polynucleotides
 CC may be used to produce the polypeptide, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the protein. The
 CC polynucleotides and its complementary sequences may also be used as DNA
 CC probes in diagnostic assays and also used in gene therapy. The
 CC and in assays to identify modulators of polypeptide expression and
 CC activity. The polypeptides and nucleic acids may be used as nutritional
 CC supplements, to modulate cytokine and cell proliferation activity, to
 CC modulate immune stimulation or suppression (e.g. for the treatment of
 CC microbial infections and autoimmune disorders such as multiple sclerosis,
 CC rheumatoid arthritis and insulin-dependent diabetes), to modulate
 CC haematopoiesis, to modulate tissue growth activity (e.g. for the
 CC treatment of Parkinson's disease, Huntington's disease and Alzheimer's
 CC disease), to modulate actin and inhibit activity (e.g. for controlling
 CC fertility), to modulate chemotactic and chemokinetic activity, to
 CC modulate haemostatic and thrombolytic activity, to modulate receptor
 CC ligand activity, to modulate inflammation and to inhibit tumour growth.
 CC
 XX Sequence 2865 BP; 488 A; 973 C; 830 G; 574 T; 0 other.

Query Match 26.0%; Score 208.8; DB 22; Length 2865;
 Best Local Similarity 63.2%; Pred. No. 2.5e-56;
 Matches 321; Conservative 0; Mismatches 187; Indels 0; Gaps 0;
 QY 1 atggccttgagagccttgagcagatgagggccttgagagattcagatcattat 60
 Db 338 atgcatcttctgacacctcctgagcctggtggtggtggtggtggtggtggtggtggtg 397
 QY 61 ctggttttattctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 120
 Db 398 acgattgctcgtgagctcctcctcctcctcctcctcctcctcctcctcctcctcctc 457
 QY 121 gctgcagccattcctgctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 180
 Db 458 tcggccgcgtgcccagcccccagcctgctggtggtggtggtggtggtggtggtggtggt 517
 QY 181 ggtatgaacttgatcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 240
 Db 518 gccagatcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 577
 QY 241 tcaatctgagggcagagagatgctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300
 Db 578 cccaaccagagagcccccagcagtgccgctcctcctcctcctcctcctcctcctcctg 637
 QY 301 ctgaatggagactatccacagcagcaagtgagggagacagaccctggtgtgtgtgtgt 360
 Db 638 cccaatgagcagggccacagcctggtggtggtggtggtggtggtggtggtggtggtg 697
 QY 361 gatatgatcaagaactcctcctcctcctcctcctcctcctcctcctcctcctcctcct 420
 Db 698 gctatgacgcagcattctcctcctcctcctcctcctcctcctcctcctcctcctcct 757
 QY 421 tatcagtaactgaatcagtggttcaatcctcctcctcctcctcctcctcctcctcctc 480
 Db 758 tctcagctctgaaagccatgagccagctcctcctcctcctcctcctcctcctcctcct 817
 QY 481 atcataagtgccatgctcctcctcctcctcctcctcctcctcctcctcctcctcctc 508
 Db 818 gctgctgagggccctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 845

PR 03-FEB-2000; 2000JP-0026862.
PR 03-MAR-2000; 2000JP-0058367.

XX (PROT-) PROTEGENE INC.
PA (SAGA) SAGAMI CHEM RES. CENT.
XX

PI Kato S, Kimura T;

DR WPI; 2001-418355/44.
DR P-PSDB; AAE06612.

XX Human proteins with hydrophobic domains and the nucleic acids encoding
PT them, useful for preventing diagnosing and treating e.g. cancer,
PT Alzheimer's and inflammation -

PS Claim 4; Page 527-532; 563pp; English.

CC The present sequence is human protein with hydrophobic domain encoding
CC CDNA clone HP03882. The polynucleotide and polypeptide of the invention
CC may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The polynucleotides
CC may be used to produce the polypeptide, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the protein. The
CC polynucleotides and its complementary sequences may also be used as DNA
CC probes in diagnostic assays and also used in gene therapy. The
CC polypeptides may also be used as antigens in the production of antibodies
CC and in assays to identify modulators of polypeptide expression and
CC activity. The polypeptides and nucleic acids may be used as nutritional
CC supplements, to modulate cytokine and cell proliferation activity, to
CC modulate immune stimulation or suppression (e.g. for the treatment of
CC microbial infections and autoimmune disorders such as multiple sclerosis,
CC rheumatoid arthritis and insulin-dependent diabetes), to modulate
CC haematopoiesis, to modulate tissue growth activity (e.g. for the
CC treatment of Parkinson's disease, Huntington's disease and Alzheimer's
CC disease), to modulate activin and inhibin activity (e.g. for controlling
CC fertility), to modulate chemotactic and chemokinetic activity, to
CC modulate haemostatic and thrombolytic activity, to modulate receptor
CC ligand activity, to modulate inflammation and to inhibit tumour growth.
XX
SQ Sequence 2194 BP; 408 A; 710 C; 615 G; 461 T; 0 other;

Query Match 22.5%; Score 180.8; DB 22; Length 2194;
Best Local Similarity 61.4%; Pred. No. 2.2e-47;
Matches 312; Conservative 0; Mismatches 187; Indels 9; Gaps 1;

QY 1 atggccttgaaggagctctgagtcgaagtggagccttggagattcagatgctcat 60
DB 58 atggcgtctcgaagccttggagcagcgagcggtggccttccagaccctgag 117
QY 61 ctggttttctctcctctcctcatttatatccctatataactgctagaaacttt 120
DB 118 gtgtcaaccttaccctccctcgtccctcattcattccagagtcctcggagaacttc 177
QY 121 gctgcagcaatcctcgtgcatcgttgcgtggtccacatgctgacaataactgactt 180
DB 178 tcagccgcacatcccgccagcagatgctggacacatgctgacaatgctctg----- 232
QY 181 ggtacgaactgggaatccctcagtgaaatgagcctcttggaactctatccactagac 240
DB 233 -----cggttccacaacagaccaccaagaccccttctgacatctccatcccgccagcgc 288
QY 241 tcaaatctgagcagcagaagatctgctgttcattccatcccaagcgagcgtcttccac 300
DB 289 cccaaccaggggcccacacagtgccgcgtcttccgacagcagatggcagcctcttgagc 348
QY 301 ctgagtggagcatatccacagcagaatgtaggacagacacccctgtgtgtagtgctg 360
DB 349 cccaatgacacagcagcagctgtagcagatgacacagcagcggtgtgtgtagcgctg 408
QY 361 gatatgatcaagactactcctcctgaccattgtgactcaagtggagcctgtagtatt 420
DB 409 gctatgacccgagcgtcttccacatcgtggtgcaagtggagcctggtgtgtagcagc 468

QY 421 tatcagtaactgaatcagtggttcaattctactctctgactggaatggtggagc 480
DB 469 tccaggggttggaagccctcagccatcccttcaactcttcaatgctcggatcgtgtgctcc 528
QY 481 atcataggtggcattgtctcagacaggt 508
DB 529 ttatctcggggcctcctcctcctacaggt 556

RESULT 15

AAE32614
ID AAE32614 standard; cDNA, 2210 BP.

AC AAE32614;

DT 24-APR-2001 (first entry)

DE Human organic anion transporter OAT4 nucleotide sequence SEQ ID NO:1.

KW Human; organic anion transporter; OAT4; nephrotrophic; kidney disease;
KW abnormal foetal growth; ss.

OS Homo sapiens.

PN WO200102562-A1.

PD 11-JAN-2001.

PF 15-JUN-2000; 2000WO-JP03878.

PR 01-JUL-1999; 99JP-0187244.

PA (NISC-) JAPAN SCT & TECHNOLOGY CORP.

PI Endou H, Sekine T, Cha SH;

DR WPI; 2001-138139/14.

DR P-PSDB; AAB69091.

PT Placental organic anion transporter gene and encoded polypeptide OAT4,
PT useful in studying causes of abnormality, and in developing drugs to
PT prevent or treat various kidney diseases and abnormal fetal growth -

PS Claim 5; Page 23-26; 32pp; Japanese.

CC The present sequence encodes a human placental organic anion transporter
CC designated OAT4. OAT4 has nephrotrophic activity. The OAT4 protein and
CC encoded gene are useful in studying causes of abnormality, including the
CC application of a variant nucleic acid as a probe to detect the presence
CC of a gene encoding the transporter OAT4, or to identify or quantify such
CC diseases and abnormal foetal growth.

SQ Sequence 2210 BP; 426 A; 709 C; 614 G; 461 T; 0 other;

Query Match 22.5%; Score 180.8; DB 22; Length 2210;
Best Local Similarity 61.4%; Pred. No. 2.3e-47;
Matches 312; Conservative 0; Mismatches 187; Indels 9; Gaps 1;

QY 1 atggccttgaaggagctctgagtcgaagtggagccttggagattcagatgctcat 60
DB 53 atggcgtctcgaagccttggagcagcgagcggtggccttccagaccctgagc 112
QY 61 ctggttttctctcctcctcatttatatccctatataactgctagagaacttt 120
DB 113 gtgtcaaccttaccctccctcgtccatgatacttccagatgctcctcggagaacttc 172
QY 121 gctgcagcaatcctcgtgcatcgttgcgtggtccacatgctgacataataactgactt 180
DB 173 tcagccgcacatcccgccagcagatgctggacacatgctgacaatgctctg----- 227


```
QY 181 gtaatagaactcgaatcctcagatgaagatgcacctcttgagaatctctatccactagac 240
Db 228 ----cgtttccacaacaacatgaaccccccaaggccctctctgacacatctccatcccgccaggc 283
QY 241 tcaaatctgagggccagagaaagtgtcgtctgttccatcccccagtgccagctctctcac 300
Db 284 cccaaccaggggccccacacagtgccgcgcttccgcgacgacacagtgccagctcttgac 343
QY 301 ctgaatgggaactatccacaagcaagtgaggcagacacagaacctgtgtgatatgcttg 360
Db 344 cccaatgccacggccacacagctggaagcgaagctgacacgagccgtgtgtgacgycgtg 403
QY 361 gtatatgatcaagctacttcccttcgaccatgtgactaagtgggaacctggtatgtgat 420
Db 404 gtcctatgaacggcagctcttccacctccaccatcgtgccaagtggaacctgtgtgacgc 463
QY 421 tatcagtcactgaaatcaatgtgttcaatccctactcttgaactggaatgctgtggaagc 480
Db 464 tcccagggtcttgaagcccccttaagccaagtcacatcttcatgttcgsggataccctgtgsgctcc 523
QY 481 atcataagtgcccatgtctcagaacaaggt 508
Db 524 ttatctgsggtccctcctcctaccaggt 551
```

Search completed: July 1, 2002, 15:45:51
Job time: 6260 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2002, 12:08:21 ; Search time 2405.18 Seconds

(without alignments)
4511.742 Million cell updates/sec

Title: US-09-674-235-10

Perfect score: 804

Sequence: 1 atggccttgagagagctctt.....acacaaagtgcacaaattt 804

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estinu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	494	61.4	615	AV652088	AV652088 AV652088
2	386	48.0	688	AG103175	AG103175 Pan trogl
3	351	43.7	399	AA680184	AA680184 z11c10.s
4	319.2	39.7	581	B1338656	B1338656 362506 MA
5	307.2	38.2	367	AA682573	AA682573 2120806.S
6	293.6	36.5	820	BG972774	BG972774 602857672
7	292.6	36.4	1054	BF783752	BF783752 602109666
8	289.6	36.0	628	BB661512	BB661512 BB661512
9	288.2	35.4	739	B1102429	B1102429 602889616
10	284.8	35.4	1914	AW107114	AW107114 um18p07.Y
11	283.2	35.0	615	BB660759	BB660759 Mus muscu
12	281.2	34.9	974	AA1530049	AA1530049 u188f01.Y
13	280.4	34.6	694	BB654082	BB654082 BB654082
14	278.2	34.2	578	AA605334	AA605334 AV605334
15	272.8	33.9	653	BB602405	BB602405 BB602405
16	272.6	33.9	754	BG566618	BG566618 602585464
17	272.6	33.9	754	BG566618	BG566618 602585464

18	266.2	33.1	628	9	BB660738	BB660738
19	265.8	33.1	807	10	B1332219	B1332219
20	264.8	32.9	492	9	AT663462	AT663462
21	261.4	32.5	608	9	AT746617	AT746617
22	261.2	32.5	741	12	AQ237446	AQ237446
23	260.4	32.4	658	9	AT1316141	AT1316141
24	255.6	31.8	637	9	BB661317	BB661317
25	253.8	31.6	714	9	AA033971	AA033971
26	253.2	31.5	505	10	B1143310	B1143310
27	253	31.5	463	9	AA259000	AA259000
28	250.6	31.2	575	9	AA475229	AA475229
29	246.2	30.6	630	9	BB661292	BB661292
30	243.4	30.3	445	9	AA475431	AA475431
31	242.6	30.2	490	9	AA106794	AA106794
32	240.4	29.9	612	9	BB637428	BB637428
33	234.2	29.1	902	10	BF785293	BF785293
34	233	29.0	811	10	B1328477	B1328477
35	230.8	28.7	451	9	AA012307	AA012307
36	230.8	28.7	612	9	AA611026	AA611026
37	230.4	28.7	915	10	BF788430	BF788430
38	229.2	28.5	607	9	AA195697	AA195697
39	229.2	28.4	570	9	AA107600	AA107600
40	226	28.1	576	9	AT1315917	AT1315917
41	214.2	26.6	730	12	BH257234	BH257234
42	207.6	25.8	617	9	BB660883	BB660883
43	200.8	25.0	571	9	BE235793	BE235793
44	198.8	24.7	512	9	AT882126	AT882126
45	197.4	24.6	901	10	B1763962	B1763962

ALIGNMENTS

RESULT 1
AV652088 615 bp mRNA linear EST 15-JAN-2002
LOCUS AV652088 GLC Homo sapiens CDNA clone GLOCWF10 3', mRNA sequence.
DEFINITION AV652088
ACCESSION AV652088
VERSION AV652088.1 GI:9873102
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 615)
XU,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.

Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)

TITLE JOURNAL MEDLINE
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1. 615
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GLOCWF10"
/clone_lib="GLC"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

Db	233	ATCCTTTTGGTCTCTTCTCCTCGTAGGCCCTTGAGTACCCATATTATACATGAGGAACCTTT	292
Qy	121	gctcagaccatctccctgcatcgctgctcgggtccacatgctcggacaataatactgtaatt	180
Db	293	ACGTGACGCCATTTCCCATCATCTCGCTCTGAGGCCCCCATCTTGCATATACATGCTCTCT	352
Qy	181	ggtaatggaactcgtgaattccctcagtcgaagtgccctcttgaaatctctatcccaactagac	240
Db	353	GATTAATGGCAGTAGGAGATCCGAGGCCAAGATGACCTCTCTAGGATCTCCATCCCCCTGGAT	412
Qy	241	tcaatcttgaggccagaagaatgctgctgcctttgtccatcccaactcggacgtcttccac	300
Db	413	TCCCACTGTGAGACTGATGATTAATGTCGTCGTTTGCCCAACCAAGTGCGCATCTTCTTCAT	472
Qy	301	ctgaaatggagactatccacagacacaagtatggagcagacacaaacctgctgtagtgctg	360
Db	473	TTGTAATGGCACTTTCTCCCAATGACACACAGACCACACACTAGCCCTGTGGATGGTTGG	532
Qy	361	gtatagatcaaaagctactcctccctcgacacatgtgactaagtggagccctgtagt	420
Db	533	GTGATGTACAGAGACCAACTTCCTTTACACATTTGTACTGCTAGTGGGACCTGGTGTGGAA	592
Qy	421	tatcatcactcgaatcaagtggttccaattccctactcttgactcgtgaatgcgcgtggagagc	480
Db	593	TCTCAGCAGCATGAATTCGTGCTCAAAATTTTCATTCATGATTTGGCTATTTATAGGGGCT	652
Qy	481	atcatagtgtagcctatgctcagacaggtg	510
Db	653	ATCATATGTGGCCATTTTGTACAGACAGTTG	682

RESULT	10
AW107114	
LOCUS	739 bp mRNA linear EST 20-OCT-1999
DEFINITION	um18h07.v1 Sugano mouse kidney mRna Mus musculus CDNA clone
IMAGE:2192701.5	similar to RF:070609.070609 PUTATIVE INTEGRAL
MEMBRANE TRANSPORT PROTEIN ; mRNA sequence.	

ORGANISM
Mus musculus
Euarctia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthalia; Rodentia; Sciurognathi; Muridae; Murine; Mus;
1 (bases 1 to 739)
REFERENCE
AUTHORS
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T.,
Underwood, K., Stepien, B., Treising, B., Allen, M., Bowers, Y., Person,
B., Skaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,
E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
TITLE
The WashU-NCI Mouse EST Project 1999
JOURNAL
Unpublished (1999)

TITLE The Mashu-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Other-ESTs: um18h07.x1
 Contact: Marra M/Mashu-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LINL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:1005153
 Seq primer: custom primer used
 High quality sequence stop: 476.
FEATURES Location/Qualifiers
 source 1..739

```
FEATURES    source
location/Qualifiers
1. .739
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:2192701"
/clone_1bp="Sugano mouse kidney mk1a"
/sex="female"
/dev_stage="adult"
```

BASE COUNT
ORIGIN

195 a 164 c 151 g 225 t 4 others

/lab.host="DH10B"
/note="Organ: kidney; Vector: pME18s-Flu3; Site_1: DraIII
(CACGCTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
(ATGGGCGCTTTTCTTTTTTTTTTTT); double-stranded cDNA was
ligated to a DraIII adaptor (TTTGGCCCTACTGG), digested
and cloned into distinct DraIII sites of the pME18s-Flu3
vector (5' site CACGCTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGCTCTAAAGCTGGC and 3' end
primer CGACCTCGAGCTCGACGACA."

Query Match	35.48;	Score 284.8;	DB 9;	Length 739;
Best Local Similarity	72.2%;	Pred. No. 8.7e63;		
Matches 367;	Conservative	0;	Mismatches 141;	Indels 0;
			Gaps	0;
OY 1 atggccttggagagcccttgatgcacagtbtgaagcccttggagattccagatgctcat 60				
Db 227 ATGTCCTTTCAGAGACTCTCTAATCAAGTTGGAGAGCTAGAGAGATCCAGATCCTTCAG 286				
OY 61 ctggcttttatcttcctctctcatgttattaaacctataccgcctaggaacttt 120				
Db 287 ATAGTTTTCTTTTCCTCCATAGCCACTTGTATACCACTATTGGAAATGGAGAACTTT 346				
OY 121 gctgcagcattctctgfgcatcagtgtgcgggtgcacagtgcgacaataaactgagct 180				
Db 347 ACTGGAGCCATTCCCAATCATGCTGCTGGGGTCCCATCCTTGACAAATGACACGCTCT 406				
OY 181 ggtaatgaactggaatccatccagttgaagatgcccctcttgaagatctatcccaactgac 240				
Db 407 GATATGAGAGTAGAGATACTGAGCCAGATGACCTCTGAGAGATCTCATGCCCCCTGGAT 466				
OY 241 tcaaatctgaagccagagaagtgtctgtgctttgctccatcccaagttgcaagctcttcac 300				
Db 467 TCCAACTGAGACTGGATTAATGTGTGCTTTTGGCCAAACACAGTGGCAATCTTTCAT 526				
OY 301 ctgaatggacatccacagcaacaagtgtggagacagacaagaacctgtgtgatatgttg 360				
Db 527 TTGAATGGCACTTCTCCAAATGTGTCAGAGCCAGACACAGACCCCTGTGTGATGTGG 586				
OY 361 gtaatgatcaaaagctacttcccttcgcacacattgtgactaagtgtggaactggtatgfat 420				
Db 587 GTGTATGACAGAGGCAAACTTCTTTCTACCAATGTGACAGTGGAGACTGGTATGTAA 646				
OY 421 taccagtcactgaataatcagttgttcaattccactcttgcagtggaatgtgtgtggagc 480				
Db 647 TCTCAGCACTGAAATCTGTCACTAAATATCAATTCATATTTGGCCTATNNATAGNGST 706				
OY 481 atcatagtggtccatgtctcagacaggt 508				
Db 707 ATCATATGTGGCCATCTGTCTAGACAGGT 734				

RESULT	11				
LOCUS	BC016496				
DEFINITION	BC016496	1914 bp	mRNA	linear	HTC 05-NOV-2001
	Mus musculus, Similar to solute carrier family 22 (organic cation transporter)-like 2, clone IMAGE:4236791, mRNA.				
ACCESSION	BC016496				
VERSION	BC016496.1	GI:16741330			
KEYWORDS	HTC,				
SOURCE	house mouse,				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 1914)				
AUTHORS	Strausberg,R.				

TITLE
JOURNAL

Direct Submission
Submitted (31-Oct-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalobcm.tmc.edu
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

REMARK
COMMENT

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 31 Row: e Column: 7
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein
This clone has the following problem: frame shifted.
Location/Qualifiers
1..1914

FEATURES
source

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4236791"
/issue_type="Kidney, normal, 5 month old male mouse."
/clone_lib="NCLCGAP_Kid14"
/lab_host="DH10b"
/note="Vector: PCMV-SPORT6"
BASE COUNT 560 a 376 c 385 g 593 t
ORIGIN

Query Match

Best local Similarity 73.6%; Score 283.2; DB 11; Length 1914;
Matches 374; Conservative 0; Mismatches 133; Indels 1; Gaps 1;

QY 1 atggccttgagagcctctgagcaagttggagccttgaggagattcagatgctcat 60
DB 156 ATGGCCTTGAGAGCCTCTGAGCAAGTTGGAGAGCTTAGAAGATTCCAGATCTTCAG 215
QY 61 ctggttttatctccctctcctcatgtattatcctcctcattactctagagaactt 120
DB 216 ATCTCTTTGTTCTCTCTCTGAGGCTTGTGTACTCATATTACATMGAGAGACTTT 275
QY 121 gctgagcattcctgctgctgagtcagtcagtcagtcagtcagtcagtcagtcagtc 180
DB 276 ACTGAGCCATTCACATCATGCTGCTGCGCCGCCATCTTGCACATGACATCGCTCT 335
QY 181 ggtatgaactgaactcctcagtaagaatgctcctctgagatctcattccactagac 240
DB 336 GATATATGCACTAGAGATCTGAGCAAGATGACCTCTGAGAGATCTCCATCCCTTGAT 395
QY 241 tcaatctgagagcagagaagtgtcgtcttctcattccctcagtcagtcagtcagtc 300
DB 396 TCCAACTGAGACTGGATTAATGTCTGTG-TTTGCCCAACCAAGTGACATCTCTTCAT 454
QY 301 ctgaatggagcattcacacacagtgaggcagacacacacacacacacacacacacac 360
DB 455 TTGAATGGCACTTCTCCAAATGAGACAGACACACACACACACACACACACACACAC 514
QY 361 gttatatacaagaactcctcctcagaccattgtgactaagtggagctggtatgtat 420
DB 515 GTGTATGACAGAGACACTTCTTCTTACCATGTGACTGATGAGACCTGCTGTGTGAA 574
QY 421 tatcagcactggaatcagtggttcaattcctactctgactgagtcgtgtggagac 480

DB 575 TCCAGGCACTGATTTCTGTGCTAAATTTTCATTCATGATGAGCCCTATTATAGGGCT 634
QY 481 atcatagtgccatgctcagacagact 508
DB 635 ATCATATGTGGCATTTGTCAGACAGCT 662

RESULT 12

LOCUS

BB660759 615 bp mRNA linear EST 26-Oct-2001
musculus cDNA clone D630003H14 5', mRNA sequence.

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES
source

BB660759 615 bp mRNA linear EST 26-Oct-2001
musculus cDNA clone D630003H14 5', mRNA sequence.
BB660759.1 GI:16494539
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 615)
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
M., Koyu, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,
Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T.,
Muramatsu, M., and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
Contact: Yoshinori Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/,
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
Wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Matahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura,
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and
Hayashizaki, Y.
RIKEN Integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multichannel sequencer. Genome Res. 10
(11), 1757-1771 (2000)
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
Y., and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Alzawa,
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., and
Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.
Location/Qualifiers
1..615
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="D630003H14"
/clone_lib="RIKEN full-length enriched, 0 day neonate
kidney"
/tissue_type="Kidney"
/dev_stage="0 day neonate"

Db 683 ATCATATGTGGC 694

RESULT 15

AV605334

LOCUS AV605334 Bos taurus kidney fetus Bos taurus linear EST 28-NOV-2001

DEFINITION 578 bp mRNA

AV605334 Bos taurus kidney fetus Bos taurus cDNA clone EIKI028B10

5', mRNA sequence.

AV605334

VERSION AV605334.1 GI:9735707

KEYWORDS EST.

SOURCE

ORGANISM

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 578)

Takasuga,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H.

and Sugimoto,Y.

Establishment of a high throughput EST sequencing system using

poly(A) tail-removed cDNA libraries and determination of 36,000

bovine ESTs

Nucleic Acids Res. 29 (22), E108 (2001)

21570554

CONTACT: Yoshikazu Sugimoto

Animal Genetics Division

Shirakawa Institute of Animal Genetics

Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan

Tel: 81-248-25-5641

Fax: 81-248-25-5725

Email: kazusugie@ccocn.ocn.ne.jp

Single pass sequencing.

This clone was obtained from a polyA-deleted cDNA library.

Location/Qualifiers

1..578

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone="EIKI028B10"

/clone_id="Bos taurus kidney fetus"

/issue_type="kidney"

/dev_stage="fetus"

/lab_host="DH10B"

/note="Vector: pZL1; Site1: SalI; Site2: NotI; Poly A

was deleted from a NotI site"

BASE COUNT 117 a 164 c 137 g 160 t

ORIGIN

Query Match 34.2%; Score 274.8; DB 9; Length 578;

Best local Similarity 74.6%; Pred. No. 3.2e-60;

Matches 367; Conservative 0; Mismatches 107; Indels 18; Gaps 1;

Db 17 tcttgagtcagtgagagccttgagagattcagatgcttcattgattatcttc 76

12 TCCCTTAATGAAGTTGGGCTGGGAAATTCAGATCCTTTCAGATGTTTTCAGCTCTTC 71

77 cctctccatgattatcatcccatatatactgctagagaaacttgctcagccattctg 136

72 CTCTGTGTGGATGACGACCCCTGATATATCTGTGGAGAAATTCACCTGCTGCATTCCTG 131

137 gtcacgttgctggttcacatgctgagacaataatactgtagtctgtaataagaaatgaa 196

132 GTGATGCTGCTGTGGTCTACATCTCTGA-----TAAATGCCACTGAGAGA 173

197 tctcagtgagagatgaccttggaatctctatcccatagactcaaatctgagagccag 256

174 TCCACAGCCCTGATGCTCTCTGAGAAATTCATCCACCTGATTCAAACTTAAAGCCAG 233

257 agaagtgctgcttgcttcacatcccccagtgagagctcttcaactgaatggaactatcc 316

234 AGAAGTGTCATGCTCTCTCCACCCAGTGGACAGCTCTTCACCTGAAATAGGACCTTCG 293

QY 317 acagcacaagtggagcagacacagaccctgtgtgtagtgcgtggtatgatgataaagct 376

Db 294 CCAACATGACACGCGCTGGACAGCGCCCTGTATGAGAGCGCTGGGTGATGACACAGAGCT 353

QY 377 actcccttgagaccatgctgtaagtgagaccctgtgtgtagtgcgtggtatgataaagct 436

Db 354 TGTTCCTCCACCATCTGCTGACCTGAGTGGACCTCATATGATGACACGAGAAAT 413

QY 437 cagtggtcattctctactctgactggaatgctggtgagagacatcattagtgagcag 496

Db 414 CAGTGTTCAGTTTGTATTCATGAGCTGTGAATGACAGTGGAGGCTTCGTATACGGGCATC 473

QY 497 tctcagacaggt 508

Db 474 TCTCAGACAGGT 485

Search completed: July 1, 2002, 14:44:48
Job time: 9387 sec


```
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F15
;
US-08-232-463-14
```

[illegible]

```

1  RESULT 5
2  US-07-906-871-15/C
3  ; Sequence 15, Application US/07906871
4  ; Patent No. 5340739
5  ;
6  ; GENERAL INFORMATION:
7  ; APPLICANT: Stevens, Richard L.
8  ; APPLICANT: Avraham, Shalom
9  ; TITLE OF INVENTION: HEMATOPOIETIC CELL SPECIFIC
10 ; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATORY ELEMENTS OF SERGLYXIN AND USES
11 ; TITLE OF INVENTION: THEREOF
12 ; NUMBER OF SEQUENCES: 18
13 ;
14 ; CORRESPONDENCE ADDRESS:
15 ; ADDRESS: Sterne, Kessler, Goldstein & Fox
16 ; STREET: 1225 Connecticut Avenue, N.W., Suite 300
17 ; CITY: Washington
18 ; STATE: DC
19 ; COUNTRY: USA
20 ; ZIP: 20036
21 ;
22 ; COMPUTER READABLE FORM:
23 ; MEDIUM TYPE: Floppy disk
24 ; COMPUTER: IBM PC compatible
25 ; OPERATING SYSTEM: PC-DOS/MS-DOS
26 ; SOFTWARE: Patentln Release #1.0, Version #1.25
27 ;
28 ; CURRENT APPLICATION DATA:
29 ; APPLICATION NUMBER: US/07/906, 871
30 ; FILING DATE: 19920103
31 ; CLASSIFICATION: 435
32 ;
33 ; PRIOR APPLICATION DATA:
34 ; APPLICATION NUMBER: US/07/816, 289
35 ; FILING DATE: 03 JAN 1992
36 ;
37 ; PRIOR APPLICATION DATA:
38 ; APPLICATION NUMBER: US/07/535, 544
39 ; FILING DATE: 18 JAN-1991
40 ;
41 ; PRIOR APPLICATION DATA:
42 ; APPLICATION NUMBER: PCT/US89/03051
43 ; FILING DATE: 13-JUL-1989
44 ;
45 ; PRIOR APPLICATION DATA:
46 ; APPLICATION NUMBER: US/07/224,035
47 ; FILING DATE: 13-JUL-1988
48 ;
49 ; ATTORNEY/AGENT INFORMATION:
50 ; NAME: Cimballa, Michele A
51 ; REGISTRATION NUMBER: 33, 851
52 ; REFERENCE/DOCKET NUMBER: 0627 .2830004
53 ;
54 ; TELECOMMUNICATION INFORMATION:
55 ; TELEPHONE: (202)833-7533
56 ; TELEFAX: (202)833-8716
57 ;
58 ; INFORMATION FOR SEQ ID NO: 15:
59 ; SEQUENCE CHARACTERISTICS:
60 ; LENGTH: 17327 base pairs
61 ; TYPE: NUCLEIC ACID
62 ; STRANDEDNESS: both
63 ; TOPOLOGY: linear
64 ;
65 ; MOLECULE TYPE: DNA
66 ;
67 ; FEATURE:
68 ; NAME/KEY: exon
69 ; LOCATION: 621..753
70 ;
71 ; FEATURE:
72 ; NAME/KEY: intron
73 ; LOCATION: 754..9596
74 ;
75 ; FEATURE:
76 ; NAME/KEY: exon
77 ; LOCATION: 9597..9744
78 ;
79 ; FEATURE:
80 ; NAME/KEY: intron
81 ; LOCATION: 9745..16396
82 ;
83 ; FEATURE:
84 ; NAME/KEY: exon
85 ; LOCATION: 16397..17327
86 ;
87 ; US-07-906-871-15

```

Query Match 4.38; Score 34.2; DB 1; Length 17327;


```
QY      788 caaagtgaccaa      801
          || | | |
Db      2143 AAACACGGAGAAA      2156
```

```

RESULT      8
US-08-796-101-39
; Sequence 39, Application US/08796101
; Patent No. 6183752
;
; GENERAL INFORMATION:
; APPLICANT: EPSTEIN, STEPHEN E.
; APPLICANT: FINKEL, TOREN
; APPLICANT: SPEIR, EDITH
; APPLICANT: ZHOU, YI FU
; APPLICANT: ZHU, JIANNUI
; APPLICANT: ERDIE, LORNE
; APPLICANT: PINCUS, STEVEN
; TITLE OF INVENTION: RESTENOSIS/ATHEROSCLEROSIS DIAGNOSIS,
; TITLE OF INVENTION: PROPTYLAXIS AND THERAPY
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/796,101
; FILING DATE: 05-FEB-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: KOWALSKI, THOMAS J.
; REGISTRATION NUMBER: 32,147
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 764-5574
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5798 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
US-08-796-101-39

```

Query Match	4.08;	Score 32.4;	DB 4;	Length 5798;
Best Local Similarity	48.0%;	Pred. 2.5;	Mismatches 131;	Indels 1;
Matches 122;	Conservative	0;	Matches 131;	Indels 1;
Qy 548 aactagatgaggcttaaaagcacttagaagaattgcagcacaatggataaagaatg				
Db 1904 AATTGGTAGACAGAAAATTTTATCTTAAACGACATACGTCTAAACAGAAATAGACATTA				
Qy 608 ctgaagaacacctgaacataagagttgtaagatccacatgcaggaagagctgtagtcag				
Db 1964 AAGAAAGTTTAGTAAACGATTAATTGCAAAATACCAATAGTTTCGAGATATATATGATCA				
Qy 668 caacgacaaacactactgtytgactgtgtcgcacaccagtatgcygtaaaagaatct				
Db 2024 TACCTACACGATGCCATATGAGATA-TTTTTTTAAACGGTACTAGAGAGAAAAGATATCT				
Qy 728 gtatccctggtatttttggagaaaaaaatctcaaggaagaaggaataaaatgatctgaca				
Db 2083 AAACCTGAGATTTTTCACGATATGATATTAAGGGAATATATTTTGGCAAAATGGGAAATA				

QY	788	caaaagtacc	801
Db	2143	AAACACGGAG	2156

```

RESULT 9
US-07-923-095-1
; Sequence 1, Application US/07923095
; Patent No. 5401652
; GENERAL INFORMATION:
; APPLICANT: Sokol Dr., Patricia T.
; APPLICANT: zhai Dr., Mohammad R.
; TITLE OF INVENTION: Nucleic Acid Sequence Encoding an Apatin
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: 1937 West Main Street
; CITY: Stamford
; STATE: CT
; COUNTRY: U.S.A.
; ZIP: 06904-0060
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,095
; FILING DATE: 19920730
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31274
; REFERENCE/DOCKET NUMBER: 31619-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 203-321-2361
; TELEFAX: 203-321-2971
; TELEX: 710-474-4059
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1730 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pig
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 140..1456
;
US-07-923-095-1

```

	Query Match	3.9%	Score 31.6;	DB 1;	Length 1730;
	Best Local Similarity	52.2%;	Pred. No. 2.4;		
	Matches 70;	Conservative	0;	Mismatches 64;	Indels 0;
				Gaps	0.
Oy	343	cctgtatgatatgctcgtggtatatagatcaagcactcccttcgacacattggtactaag	402		
Db	1529	ccctttatgctgggctttgtttgtttttaaagccacgcgaatggcacaacttactctcag	1588		
Oy	403	tgggaacctgcatgtgattatcatcagtcactgaatcagtggttcaatcctactctcgtac	462		
Db	1589	tgggaagatgcacaatgagatctaccaggggtgttgaaataatttctgaattttccacch	1648		
Oy	463	ggaatgcctggtggtg	476		
Db	1649	tgaatgctgagctg	1662		

APPLICANT: Zlati Dr., Mohammad R.
TITLE OF INVENTION: Nucleic Acid Sequence Encoding an Apamin
Patent No. 5652111
TITLE OF INVENTION: Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: 1937 West Main Street
CITY: Stamford
STATE: CT
COUNTRY: U.S.A.
ZIP: 06904-0060
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,716
FILING DATE: 08-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/923,095
FILING DATE: 30-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lomey Dr., Karen A.
REGISTRATION NUMBER: 31274
REFERENCE/DOCKET NUMBER: 31619-00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 203-321-2361
TELEFAX: 203-321-2971
TELEX: 710-474-4059
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1730 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pig
FEATURE:
NAME/KEY: CDS
LOCATION: 140..1456
US-08-436-716-1

Query Match 3.9%; Score 31.6; DB 1; Length 1730;
Best Local Similarity 52.2%; Pred. No. 2.4;
Matches 70; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 343 cccgtgtgatgctggtatataatgaagctactccctcgacatgtgactaaag 402
DB 1529 ccccttatatgctggcctgtttgttttttaaacgacgacgaacatgacacgtcag 1588
QY 403 tgggaacctgtatgtatatacagtcactgaatcagtggttaactcactactcgtact 462
DB 1589 tgggaacttgcgaatagatgacgacgaggggtgtttagaataattctgaattttccacct 1648
QY 463 ggaatgctgtgtgg 476
DB 1649 tgaatgctgagtg 1662

RESULT 13
PCT-US92-06840-1/c
Sequence 1, Application PC/TUS9206840
GENERAL INFORMATION:
APPLICANT: Shi, Yang
APPLICANT: Seto, Edward
APPLICANT: Shenk, Thomas

TITLE OF INVENTION: YX1 TRANSCRIPTION FACTOR AND METHODS OF
TITLE OF INVENTION: ISOLATING SAME
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
STREET: 1180 Avenue of the Americas - 7th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-8403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06840
FILING DATE: 19920814
CLASSIFICATION: AU 1805
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/746,485
FILING DATE: 16-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Dennis, Manette
REGISTRATION NUMBER: 30,623
REFERENCE/DOCKET NUMBER: M-12594 CIP (1570-8)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 382-0700
TELEFAX: (212) 382-0888
TELEX: 236925
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2353 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Hela cells derived from cervical
CELL TYPE: tumor cells
CELL LINE: Hela
IMMEDIATE SOURCE:
LIBRARY: D98/AH-2
CLONE: p14-1 or pY1
FEATURE:
NAME/KEY: CDS
LOCATION: 241..1485
PCT-US92-06840-1

Query Match 3.9%; Score 31.4; DB 5; Length 2353;
Best Local Similarity 52.7%; Pred. No. 3.3;
Matches 68; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 669 acagaccaaactactctgtgtgactgttcgcgaaccccaagtggttaagagactg 728
DB 2273 ACAAGTCAGACACACACTTCTGTAAGTGAACAGCAAGAAATTGAACTGAACATCAGCAT 2214
QY 729 tatcctgtatattttgagaataaatactcaaggaagaagcataaataatgtgtacac 788
DB 2213 CTGGCAGTATTTTGTGAAAAAAGTACTAATAATGGCTTAATTTGATTAACTA 2154
QY 789 aaagtac 797
DB 2153 TTAATACAC 2145

RESULT 14

US-09-541-782-9
; Sequence 9, Application US/09541782
; Patent No. 6284480
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/541,782
; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3741
; TYPE: DNA
; ORGANISM: H.sapiens
US-09-541-782-9

Query Match 3.9%; Score 31.4; DB 4; Length 3741;
Best Local Similarity 57.7%; Pred. No. 4.2; Mismatches 41; Indels 0; Gaps 0;
Matches 56; Conservative 0;

OY 589 acaaatggaataaagaatgcggaagaacccctgaacatagaggtgtgaagatccaccatg 648
Db 1193 aaaaatggagtgatattctcgaagaatattagatcatgagtgagaaatgaactgtt 1252
OY 649 cagagagagctgctgacgacacccaactactg 685
Db 1253 caagaagagcagatgtgtaattgtgaaaaaatgt 1289

RESULT 15
US-09-141-000-4/C
; Sequence 4, Application US/09141000
; Patent No. 6054295
; GENERAL INFORMATION:
; APPLICANT: Chen, Fang
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
; FILE REFERENCE: 199999Y
; CURRENT APPLICATION NUMBER: US/09/141,000
; CURRENT FILING DATE: 1998-08-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Human
US-09-141-000-4

Query Match 3.9%; Score 31; DB 3; Length 458;
Best Local Similarity 9.2%; Pred. No. 1.8;
Matches 31; Conservative 102; Mismatches 205; Indels 0; Gaps 0;

OY 35 gccctggagatccagatcctcatcgttttattcttcctcctcatgttataa 94
Db 444 BMC.M.NR.D.BTMSA.Y...AK.KMCTYY.H.KD.CT.RH..T.D.BH..M.BT.B 385
OY 95 tccctataactactagagaacttgctgacacatctcctgcatcgttgcgtggtcc 154
Db 384 H..DKSHSNT.T.TM.AB...M..MKSMRMB...TNN.H..CT.MS.H.HK.RHHTRB 325
OY 155 acatgctgacataataactggtatgaactgaatcctcagatgaagatgccc 214
Db 324 ..H..S.SYRBC...KWTs...SK.HT.S.AS.C..DMTWC..BB..YHT.HG.AA. 265
OY 215 tcttgaatcctatccactagactcaatctgagccagagaagtgcgtgcttgg 274
Db 264 TM.HSH.BA.H.KTR.M...TB..DSB.MNR.MMT..B...N.R.SN.TH.YYMRM.YCC 205

OY 275 tccatcccccagtgagccttcttcaactgaatgggactatccacagcacagtgaagcag 334
Db 204 YBYH.YBC.M.KCBM.GMK.YGT.GSMYYMA..G.NAT.GSR..NC..AYM..TWGT.G 145
OY 335 acacagacccctgtgtgagtggtggtatgatgacaa 372
Db 144 STBCRDRCST.HCGB.G.YM.KSN.KR.GMA..H.B.A 107

Search completed: July 1, 2002, 15:40:20
Job time: 10469 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2002, 14:44:48 ; Search time 2405.18 Seconds
(without alignments)
9567.811 Million cell updates/sec

Title: US-09-674-235-19
Perfect score: 1705
Sequence: 1 aagacgcgaggaagctctt.....aatttaaaattatattc 1705

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: em_estha:*
2: em_estha:*
3: em_estha:*
4: em_estha:*
5: em_estha:*
6: em_estha:*
7: em_estha:*
8: em_estha:*
9: gb_est1:*
10: gb_est2:*
11: gb_hnc:*
12: gb_hnc:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	549	32.2	615	9	AV652088
2	545	32.0	779	9	AV652088
3	388	22.8	688	12	AG103175 Pan trolg
4	387	22.7	399	9	AA680184
5	382	22.6	612	12	AO280703 CITR1-EI
6	372.8	21.9	474	12	AO127694 HS_3091_B
7	370	21.7	578	12	AO280563 CITR1-EI
8	346.2	20.3	367	9	AA682573 z120a06.s
9	336.4	19.7	754	10	AA682573 z120a06.s
10	328.8	19.3	581	10	AA682573 z120a06.s
11	314	18.4	741	12	AO237446 RPK111-69
12	298.2	17.5	820	10	AO237446 RPK111-69
13	293.8	17.2	866	10	BI102429 60289616
14	292.6	17.2	739	9	AM107114 um18h07.Y
15	289.8	17.0	1054	10	BF783752 602109666
16	289.8	17.0	628	9	BB661512
17	287.8	16.9	1914	11	BC016496 Mus muscu

18	284	16.7	974	9	AI330049	AI330049 um8f01.Y
19	282.8	16.6	694	9	BB654082	BB654082 BB654082
20	281.4	16.5	615	9	BB660759	BB660759 BB660759
21	280.6	16.5	653	9	BB602405	BB602405 BB602405
22	274.8	16.1	578	9	AV605334	AV605334 AV605334
23	270.8	15.9	628	9	BB660738	BB660738 BB660738
24	269.2	15.8	608	9	AI746617	AI746617 um10d05.Y
25	268.6	15.8	658	9	AI316141	AI316141 u125e04.Y
26	268.2	15.7	901	10	BI763962	BI763962 603049842
27	266.8	15.6	807	10	BI332219	BI332219 602981678
28	265	15.5	492	9	AI663462	AI663462 UK3305.Y
29	260.2	15.3	637	9	BB661317	BB661317 BB661317
30	253.8	14.9	714	9	AA033971	AA033971 z105c09.Y
31	253.4	14.9	505	10	BI143310	BI143310 602907858
32	253	14.8	463	9	AA259000	AA259000 um63f07.Y
33	250.8	14.7	575	9	AA475229	AA475229 um63f07.Y
34	250.8	14.7	630	9	BB661292	BB661292 um63f07.Y
35	247	14.5	915	10	BE788430	BE788430 602114177
36	245	14.4	612	9	BB637428	BB637428 BB637428
37	243.4	14.3	445	9	AA475431	AA475431 um66c08.Y
38	242.8	14.2	490	9	AA106794	AA106794 um33f07.Y
39	238.8	14.0	902	10	BE785293	BE785293 602108493
40	238.6	14.0	612	9	AA611026	AA611026 um75f02.Y
41	237.6	13.9	811	10	BI328477	BI328477 602985750
42	237.6	13.9	607	9	AI956697	AI956697 u175d06.Y
43	236.4	13.9	570	9	AA107600	AA107600 u193a06.Y
44	233.8	13.7	576	9	AI315917	AI315917 u127f05.Y
45	230.8	13.5	451	9	AA012307	AA012307 um07c06.Y

ALIGNMENTS

RESULT 1
AV652088 615 bp mRNA linear EST 15-JAN-2002
LOCUS AV652088 GLC Homo sapiens cDNA clone GLCWF10 3', mRNA sequence.
DEFINITION AV652088
ACCESSION AV652088
VERSION AV652088.1 GI:9873102

KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 615)
Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.

Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)

JOURNAL
MEDLINE
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn

FEATURES
source
Location/Qualifiers
1..615
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GLCWF10"
/clone.lib="GLC"
/tissue.type="Corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 146 a 139 c 148 g 180 t 2 others
ORIGIN

Query Match 32.2%; Score 549; DB 9; Length 615;
Best Local Similarity 91.8%; Pred. No. 7.7e-82;
Matches 613; Conservative 0; Mismatches 1; Indels 54; Gaps 1;

```

QY 14 gctctccactacagcgtcttattgctggtggtccggagccatgagatgaatgat 73
DB 1 GCTCTTCCACACAGCGCTGTATTGCACTGCTGAGTCCGGCCCATGAGAAATGAT 60
QY 74 gggagagatacaacaagcttaattgaataaagaagaatattctcccttgaac 133
DB 61 GCGAGATCAATACAAAGCTTAATTGAATTAAT-----
QY 134 ttatctccglaagcaatgctgctccctcttgagggtcagtggttcaaatcagct 193
DB 94 -----TCTTGGGGGTCACTGTTCCAAATCAATGGCT 126
QY 194 ttgagagactcttgagcaagttgagagcccttgagagattcagatgctcattgtt 253
DB 127 TTGAGGACCTCTGAGTCAAGTTGAGGACCTTGAGATTCAGATGCTTCATCTG 186
QY 254 ttattctccctctcctatgattatataatccctacatactcgtagaagacttgcag 313
DB 187 TTATCTTCCCTCTCTCTCATGTTAATATCCCATATCTGCTAGAGAACTTGTG 246
QY 314 ccaatctcgtgcatcgttgcgtgggtccacatgctcgagacaataatcagctgta 373
DB 247 CCATTCCTGTCATCGTGTGCTGGTCCACATGCTGACAAATATCTGATCTGTA 306
QY 374 aaactggaatcctcagagaagatgctcctcttgagaatctcctacactagaatc 433
DB 307 AAATGGAATTCCTCAGTGAAGATGCCCTTGAAGATCTATCCACTAGACTCAATC 366
QY 434 tggagcagaagatgctgctgcttgcacatccctgagctgctcctcctcagatg 493
DB 367 TGAAGCCAGAGAGTGCCTGCTTGTTCATCCCGAGGACCTCTTACCTGTAATG 426
QY 494 ggcataccacagacagaagagagagacagaaacctgtgtcagatgctggtatag 553
DB 427 GGCATTCACACAGCAGAGAGGACACACAAACCTGTGTGATGGCTGGGTATAG 486
QY 554 atcaagaactactcctcctgacacattgactgaagtgagacctgagtatcagat 613
DB 487 ATCAAGAAGCTACTCTCTTGCACATGTGACTAAGTGGACCTGGTATGATATCAGT 546
QY 614 cactgaataatcagtttcaatcctcctcctcctcctcctcctcctcctcctc 673
DB 547 CACTGAATACAGTGTGATCTTACTTCTGACTGAGATGCTGGTGGAGCATATAN 606
QY 674 gtggccat 681
DB 607 GTGGCCAT 614

```

RESULT 2
AV699858/c 779 bp mRNA linear EST 16-JAN-2002
LOCUS AV699858 GKC Homo sapiens cDNA clone GKCDEL04 3', mRNA sequence.
DEFINITION AV699858
ACCESSION AV699858
VERSION AV699858.1 GI:10301829
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 779)

Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.

TITLE

Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106

COMMENT

Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@hgsc.sh.cn

FEATURES

This clone is available at CHGC in Shanghai.
Location/Qualifiers

1..779

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="GKDEL04"

/clone_lib="GKC"

/tissue_type="hepatocellular carcinoma"

/dev_stage="Adult"

/lab_host="SOLR"

/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 250 a 122 c 110 g 292 t 5 others
ORIGIN

Query Match 32.0%; Score 545; DB 9; Length 779;
Best Local Similarity 91.6%; Pred. No. 3.3e-81;
Matches 619; Conservative 0; Mismatches 50; Indels 7; Gaps 4;

```

QY 1030 aaaaaaatacagaagaacacacacagaaggttttccctacacacagaagac 1089
DB 703 AGAAAAAATTCAGAGAAAGAACCCNCCAG--GGGTTTTTCTTCAACCCAGCAAAAC 646
QY 1090 atatatagatactgatactcactaattatcagatcattatcattatc 1149
DB 645 ATATTTTAAACACATGATCCATTAATTAATAT--GGCATTAATTCATTTATTTCANA 590
QY 1150 attaaacttgaggacatgatactcctcctgagcaatcagatatttggagaatccttaa 1209
DB 589 ATTACNTTGTGGGACATGTAATCT--TTGAGCAATCTGATATTTTGGGAAAGTCTTTAA 531
QY 1210 aaagttaacaaattatcaaaaattactagtagaataagatgattcagaacaaaga 1269
DB 530 AAAGTTCCAAATTTATTCATAATTTCTAGTAGATTGAAGATTGAGAACCAAGAAAGAAA 471
QY 1270 tcacagaattagatgtgctgctgctgtgtatgaagcacatgtgtatgaattc 1329
DB 470 TCCAGAAATTAGGATGTGCTGCTGCTGTATGACACCACTGTGATGAAATTCATTAAGT 411
QY 1330 tgcacaaagtcaaaacatactgttcaatgcaacacagaatacaataaataacaga 1389
DB 410 TCCAAAAGTCAAAACCAATCTGTACATGCACACAGAAATCAAAATTAATTCAGAAATAGA 351
QY 1390 gacctataaataatgataatcactgatacttttgacataataagccattgga 1449
DB 350 GACCTATATTAATGATTAATACATGATCTTTGACATTAATTAAGCCATTGGAAAAAGG 291
QY 1450 aaagttagatacctaataaactgatactccttctgtaataacacagatcactaa 1509
DB 290 AAAGATTGATATCAAAATTAATACATGATCTTTGTAATTAATTAATTAATTAAT 231
QY 1510 tagtactctccatggtggaatttaattccttttcttctgtatatttctctgtat 1569
DB 230 TAGTTACTTTCCATGGGGAATTTTAATTAATCTTTCTTTGTAATTTTCTCTGTAT 171
QY 1570 atttaacaaatagctgtgataagttacaaatataaagaatattgtaacatgaag 1629
DB 170 ATTTTAAACAAATAGCTGATATGATTTCAATATTAATTAATTAATTAATTAATTAAGG 111
QY 1630 gcaaaagccagttcagcaatttcaaacctgtatgatacatttaataaataacta 1689

```


	Db	110	GCAAGCCGCTTCACACTGTATGTACATTAAATAAATACCTTAAT	51
Oy	1690	taaaactatatttc	1705	
Db	50	TAAAAATTATTTCTC	35	

RESULT 3							
LOCUS	AGI03175	688 bp	DNA	linear	GSS 03-NOV-2001		
DEFINITION	Pan troglodytes DNA, clone: PTB-106L08.F, genomic survey sequence.						
ACCSSION	AGI03175						
VERSION	AGI03175.1 GI:16723692						
KEYWORDS	GSS; GSS (genome survey sequence).						
SOURCE	Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male BAC library clone:PTB-106L08.F.						
ORGANISM	Pan troglodytes						
REFERENCE AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.						
TITLE	1 (sites)						
JOURNAL REFERENCE	Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Tokokl,Y., Watanabe,H. and Sakaki,Y.						
AUTHORS	BAC end sequences of library PTB						
TITLE	Unpublished						
REFERENCE	2 (bases 1 to 688)						
AUTHORS	Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Tokokl,Y., Watanabe,H. and Sakaki,Y.						
TITLE	Direct Submission						
JOURNAL	Submitted (02-AUG-2001) Aao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Shuliro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimbessgsc.riken.go.jp, URL:http://hpq.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)						
COMMENT	C clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the Red process and may have higher chance of clone tracking errors.						
	PRIMERS						
	Sequencing: -21M13						
	LIBRARY						
	Vector : PKS145						
	R.Site 1 : SacI						
	R.Site 2 : Sacc.						
	Location/Qualifiers						
FEATURES	/organism="Pan troglodytes"						
source	/db_xref="taxon:9598"						
	/clone="PTB-106L08.F"						
	/sex="male"						
	/cell_type="lymphoblast"						
	/clone_lib="PTB Chimpanzee Male BAC Library"						
BASE COUNT	162 a 160 c 154 g 207 t						
ORIGIN							
	Quality Match	22.8%;	Score 388;	DB 12;	Length 688;		
	Best Local Similarity	96.4%;	Pred. No. 4..3e-95;				
	Matches	397;	Conservative	0;	Mismatches	15;	Indels
Oy	177	ttacaacaaatggccttgcaggaccttggaacgttgcagaagtgcgaagcccttgsgaatltca	236				
Db	43	tttcgcgacgrtcgacgcttcgccgacgctttagtcacaaattcgagcccttcggagatttca	102				
Oy	237	gatgtcatcatcgtgttttatctctccctctcacatgtatlaattcaatcccatatacgtc	296				
Db	103	gattcctaatactgggttttttatcttcctccctctcatgtatttaattcccctaataactgct	162				
Oy	297	agagaacttgtctgcagccattccctgctcatcatgttcgtgcgttccaacatcgtgcgaataa	356				
Db	163	AGAACAATTGGTGCAGGCCATTCTCGTGATCATGCTTGCTGGGTCACATGTCGAACATAA	222				
Oy	357	tactgatcttgttaatagaactggaatccctcagtygaagatgcctcttgagaatcctcat	416				

Db	223	TACGATCTGTAATGAAACCTGGAATCCTCAGTGAAGATGCCCTTGAAATTTCTAT	282
OY	417	ccactagatcaaatctgaggccagaagaagtgtcgtcgtcttcttgcattcccccagttgca	476
Db	283	CCGACTGACTCAAAATCTGAGGCCACAGAAAGTGCTGTGCTTTGTCATCCCCAGTGGCA	342
OY	477	gctcttcacctgaataggagactatccacagacaaagtggagagacacagaacctgtgt	536
Db	343	GCTTCTTCACTGAAAGGACTATCCACAGCACAAGATGAGGACACAGAAACCTGTGT	402
OY	537	ggatgctgggtatgatgcaaaagtactctcccttgaccattgactgaag	588
Db	403	GGAGGCGTGAGTATGATGATCAAGACTCTCCCTTGACACATTGTGACTAAG	454
RESULT	4		
AA680184/c		399 bp	mrna linear EST 19-DEC-1997
LOCUS	AA680184		
DEFINITION	z11c10.s1 Soares_fetal_liver_spleen.lnFLS.S1 Homo sapiens CDNA		
ACCESSION	AA680184		
VERSION	AA680184.1	GI:2656651	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 399)		
AUTHORS	Hillier, L., Allen, M., Bowles, J., Dubnue, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Maitra, M., Marín, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, T., Wyllie, T., Waterston, R. and Wilson, R.		
TITLE	Washington University School of Medicine		
JOURNAL	Washington University School of Medicine		
COMMENT	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: estw@wustl.edu This clone is available royalty-free through LIND, contact the IMAGE Consortium (info@image.jnl.gov) for further information. Seq primer: -40m13 fwd. ET from Amersham. Location/Qualifiers 1. 399 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:430482" /clone_lib="Soares_fetal_liver_spleen.lnFLS.S1" /sex="male" /dev_stage="20 week-post conception fetus" /lib_host="DH10B (ampicillin resistant)" /note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; This is a subtracted version of the original Soares fetal liver spleen lnFLS library. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5', AAGCGAAGAATTAATTAAGATCTTTTCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaído."		
FEATURES			
SOURCE			
BASE COUNT	95 a 100 c	72 g 132 t	
ORIGIN			
Query Match	22.7%	Score 387;	DB 9; Length 399;
Best Local Similarity	99.7%	Pred. No. 7.3e-55;	
Matches 398; Conservative	0;	Mismatches 0;	Indels 1; Gaps 1.
629	ttcaatcttactctgactgagtaatgctgtgtgaggagcagcatataggtgagcattgtccag	688	

	Query Match	22.4%	Score 382	DB 12	Length 612	
	Best Local Similarity	100.0%	Pred. No. 4.4e-54			
	Matches 382	Conservative	0	Mismatches	0	Indels
					0	Gaps
QY	1324	taaeagtgcacaagctcaacaacatctgctgcacgcacccagaatcaataatccaga	1383			
Db	1	TAAAGTTCGCAAAAGTCAAAACATCTGTCATGCAACCAAGAAATCAAAATTAATCCAGA	60			
QY	1384	aatagagacctatataatgaatcattatcatgatattctgacataatgaacattgga	1443			
Db	61	AAATGAGACCTATATTAATGCATTTATCATGATCATCTTTGACATTAATAAAGCCATTGGA	120			
QY	1444	aaacggaaaagattagatgactactaataacatctgctatctcttggtaataacagtcactaa	1503			
Db	121	AAACGAAAGATTAGTACGTAAATAATACATGTGACTATCTTGTGTAATATACGTCACCTAAA	180			
QY	1504	tgaagttagtactcttccatctgctggaatttcaattactcttctcttgaattttctctc	1563			
Db	181	TGATGTTAGTACTTTTCCATCGTGGAAATTTAATTACTTTTCTTGTAAATTTTTCCT	240			
QY	1564	cgtatattttaacaacaataagctggtatagtttacaatatataaagaatctgtccaat	1623			
Db	241	CTGTATATTTTAAACAAATAGCTGGTATAGTTTACAAATATATATAAGATATTGTTCAAT	300			

RESULT	6
AQ127694	
LOCUS	
A0127694	

ACCESSION	sapiens genomic clone Plate=3091 Col=4 Row=D,	DNA sequence.
VERSION	AQ127694	
	AQ127694.1	GT..350A6C0

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 474)

TITLE	JOURNAL MEDLINE	COMMENT
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome	Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)	
Contact: Mahatras GG, Wallace JC, Hood L	High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618	

```

FEATURES
  source
    fax: (206) 616-3887
    Email: jwallace@u.washington.edu
    Sequence Tagged Connector
    Plate: 3091 Row: D Column: 4
    Class: BAC ends
    High quality sequence stop: 474.
    Location/Qualifiers
      1..474
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="Plate=3091 Col=4 Row=D"

```


Query Match	Best Local Similarity	Score	DB	Length
Matches 387; Conservative	72.3%;	298.2;	10;	820;
		Pred. No. 3.4e-40;		
		Mismatches 148; Indels		Gaps
160	ctcttgagggtacgctgttacaacaataaaggcccttgagagctcttgtaagtga	219		
152	cccatggagaaagtcattcacttccaatggcccttcagaaactcctgaataggga	211		
220	ggcccttgagagattcagatgtcttactcgtgtttattctccctctcagttaata	279		
212	agccttagaagaattccacagcttcttagagatctctttgttctctgtagcccttg	271		
280	atccctatactctgctagagaacttgcctgagccatccctgctatcgtctggtgc	339		
272	gtactctaatattacatggagaaacttctgagccatttcccaatcactgctgccc	331		
340	cacatgctgagacaataactactgagatctgtaataaacttgaaatcctcagtgaatg	399		
332	cccatctcttgaaatgaacactgctcttattatggcactgagatcctgagcaagatgac	391		
400	ctcttgagaatctctatactccactgactcaaatctgagccagagaagtgcgtctt	459		
392	ctcttgagagatctccatccctctgatttccaaactgagactgattatgctgctgtt	451		
460	gtcatccccaagtgcagactctctcactgaatgagacttccacagcaaatgagga	519		
452	gcccaacacagtgccgactctcttcttcttgaatggcacttcttccaatgagacagcca	511		
520	gacacaaacccgtgtgtagctgagctggtgataatgaataaagctactctccctgacatt	579		
512	gacacttagaccctgctgtagtggtgggtatgacagagcaactcttcttaccatt	571		
580	gtgactaagtgagaccgtgtagtattatcagatcactgaaatcagatggttcaactcta	639		
572	gtgactagtgagaccgtgtagtattatcagatcactgaaatcagatggttcaactcta	631		
640	ctcttgactggaatgctgtgtgaggagcattcagtggtgcattctccacagatg	694		
632	ttcatgattggcctattttataggggctatcatattgaggcattttgacacagct	686		

RESULT 13

LOCUS B1102429

DEFINITION 602889616r1 NCI-CGAP_Kid14 Mus musculus cDNA IMAGE:5044850

ACCESSION B1102429

VERSION B1102429.1

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 866)

NIH-MGC <http://mgc.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg Ph.D

866 bp

mRNA

linear

EST 26-JUN-2001

IMAGE:5044850

QY 400 ctcttgagaaactctatccactagactcaaatctgagcgaagaagtgtcgtcgtt 459
 Db 440 CTCTTGAGGATCTTCATCCCTCCCTGATTTCCAACTGAGAGATTAATGTCGCTTTT 499
 QY 460 gtccatccccaagtgagactcttcttaacctgaatgggaactccaagacaagtggagca 519
 Db 500 GCCCAACACAGAGGACCTCTTCATTTGAATGGACCTTTCACAAATGTCAGAGACCA 559
 QY 520 gacacgaacccctgtgtgagctgtgataatgaatgaatctcccttgagacatt 579
 Db 560 GACACTGAGCCCTGTGTGATGCTTGGCTATGACAGGACCAACTTCTTCTACCATTT 619
 QY 580 gtgactaagtggagacctgtatgtatcatcaagtcacatgaatcagtggtcaattccta 639
 Db 620 GTGACAGAGTGGGACTTGTATGTGATCTCANGCACTGAATTTCTGACTAATATATCA 679
 QY 640 ctctcagcagaaatgtctgtggagcattcatagtggtgcattctcagacaggt 694
 Db 680 TTCATGATTGGCCTTATNTATAGNGSGATATCATATGTGGCATCTGTACAGACAGT 734

RESULT 15

BF783752

LOCUS 1054 bp mRNA linear EST 12-Jan-2001

DEFINITION 602109666F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4237793

ACCESSION BF783752

VERSION BF783752.1 GI:12088788

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1054)

NIH-MGC http://mgc.ncl.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

CDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

Clone Sequencing by: Incyte Genomics, Inc.

found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov

Plate: LLM9848 row: d column: 18

High quality sequence stop: 665.

Location/Qualifiers

1..1054

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:4237793"

/clone_lid="NCI CGAP_Kid14"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: Kidney; Vector: PCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.75 kb. Constructed by Life

Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 276 a 250 c 256 g 272 t

ORIGIN

Query Match 17.2%; Score 292.6; DB 10; Length 1054;
 Best Local Similarity 73.6%; Pred. No. 2,7e-39;
 Matches 373; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 188 tggcccttgaggagctcttaagttaaggagccttgagagattcaagatgtctcattc 247
 Db 1 TGGCCTTTCAGGACCTCTGAATCAAGTTGGAACCTTAGGAAAGATTCAGATCCTTCAGA 60

* QY 248 tggttttattctccctctcatgtattatccctcatatctgtctagagaacttg 307

Db 61 TCTTTTGTGTTCTTCTCTCTAGCCCTTGGTGATCTCATATTAACATGGAGAACTTTA 120
 QY 308 ctgcagccattcctctgcatcgtgtctgtggtccacatctgagacaataatcgtatctg 367
 Db 121 CTCGAGCCATTTCCCAATCATGCTGCTGGGCCCCCATCTTGACAAATGACACTGCTTG 180
 QY 368 gtaatgaactgtgaatccatcgaatgaatgacctcttgagaatctctatcccaatgact 427
 Db 181 ATATGGCACTAGAGATCTCTGAGCCCAAGATGACCTCTGAGAGATCTCATGCCCCGTGATT 240
 QY 428 caaatctgagagccagagaagtggtcgtctgttccatcccaagtgtgcaactcttaacc 487
 Db 241 CCAACCTGAGACTGGATTAATGTGCTGTTTGCCCAACCAAGATGGCACTTCTTCATT 300
 QY 488 tgaatggactatccacagacaagtgagagcagacagaaacctgtgtgagtggtggtg 547
 Db 301 TGAATGGCACTTCTCCAAATGAGACAGAGCCAGACACTGAGCCCTGTGTGATGTTGGG 360
 QY 548 tataatgaactgaactcctctcgaacctgtgtgaactaagtgtgacctgtgtatgtatt 607
 Db 361 TGTATGACAGAGCACTTCTCTTCTTACCAATGTGACTGAGTGGACCTGTGTGTAAT 420
 QY 608 atcagtcactgaatcgaatgaatgttcaatctcactctctgaactgaatgctgtgtggagca 667
 Db 421 CTCAGGCACCTGAATTTCTGTGTTCAATTTTTCATTTTCATGATTTGCTTATTTAGGGGCTA 480
 QY 668 tcaatggtgtgcattgtctcagacaggt 694
 Db 481 TCATATGTGGCCATTTGTGTCAGACAGGT 507

Search completed: July 1, 2002, 14:44:55
 J00 time: 9394 sec

RESULT 2
US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:

[illegible]

```

RESULT 3
US-07-867-106-2
; Sequence 2, Application US/07867106
; Patent No. 5389526
;
; GENERAL INFORMATION:
;
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Ketch L
; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
; NUMBER OF SEQUENCES: 19
;
; CORRESPONDENCE ADDRESS:
; ADDRESS: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: PA
;
; COUNTRY: USA
; ZIP: 19103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/867, 106
; FILING DATE: 19920625
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PJ 7187
; APPLICATION NUMBER: PCT/AU90/00530
; FILING DATE: 02-NOV-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Feeney, Joanne Longo
; REGISTRATION NUMBER: 35,134
; REFERENCE/DOCKET NUMBER: RICE-0002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5852 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
; NAME/KEY: CDS
; LOCATION: 2378..5038
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
;
; US-07-867-106-2

```

Query Match	3.2%;	Score 54;	DB 1;	Length 5852;
Best Local Similarity	49.1%;	Pred. No. 0.00025;		
Matches 173;	Conservative	0;	Mismatches 175;	T-33

QY 1345 aatcgtctacatgcacccgaatacaataaataatccagaatagagccatataatgc 1404
Db 5331 AATTATTTGATTTTAAAAAATAAAAAAATAAAAAAATAATCAATATGTTTATG 5490
QY 1405 attaatacatgatactttgacataaagccattggaacagaaagattgatacca 1464
||||| ||||| ||||| |||||

QY	479	tcttcacactgaatgggactataccacgacaaagtgaagcagacacgaaccctgtgtg	538
Db	349	TGCATCTGCCAAACCCAGTCTTCCCAATACACCCAGGGGGCCACCGAGGCATGCTTGG	408
QY	539	atggcctgggtatgatcatcaagctacttccttcgcacatttgcactaagtgggaccttg	598
Db	409	ATGGCTGGATCT-----ACAAACGACACGAGACACCATTTGACAGAGTGGACTTGG	467
QY	599	tatgtattatcagtcactgaatcaatcagttgttcaattccctactctgcagtgaacgtcg	658
Db	463	TATGCCGCTCCAAACAACGATGAAGAGATGCGACAGTCAAGCTTCATGCGAGGTATACGG	522
QY	659	tgggaagcatcatagtgccatgtcttaagcaagt	694
Db	523	TTGGAGACCTGTGTTGGAGAACTGTAGACAGGT	558

RESULT 5
 US-08-487-826B-13
 Sequence 13, Application US/08487826B
 Patent No. 5993827
 GENERAL INFORMATION:
 APPLICANT: Sim, Kim L.
 APPLICANT: Chitnis, Chetan
 APPLICANT: Miller, Louis H.
 APPLICANT: Peterson, David S.
 APPLICANT: Su, Xin-zhuan
 APPLICANT: Williams, Thomas E.
 TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
 TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US
 ZIP: 92660
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/487,826B
 FILING DATE: 10-SEP-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Israelson, Ned
 REGISTRATION NUMBER: 29,655
 REFERENCE/DOCKET NUMBER: NIH121.001CPI
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 19124 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-487-826B-13

		2.9%	Score 49.6;	DB 2;	Length 19124;	
Query Match			Pred. No. 0.0054;			
Best Local Similarity	43.8%		Mismatches 279;	Indels 0;	Gaps 0;	
Matches 217;	Conservative					
QY 1021	aactcagaanaaaatcacggaagacacaccggagtcttcctcaaaccc	1080				
Db 15481	AATTTTTAAAAATAAAAAAAAAAAAAAAAAGAGAAAATTTTTTAAAAAATAATA	15540				

; ORIGINAL SOURCE:
 ; INDIVIDUAL ISOLATE: 23-5
 US-08-731-722-5

Query Match	2.9%	Score 48.8;	DB 2;	Length 1186;
Best Local Similarly	45.7%	Pred. No. 0.0027;		
Matches 290; Conservative	0;	Mismatches 327;		

[illegible][illegible][illegible]

1304 agccacatgcatgaattcaataaagttgcacaaagttcaacaacatctgtacatgcaacca 1363
 |||||
 401 CATTTATGTAATTAATTAATCTTTTAAA--CATATATAACTTTTAAAAAAATTT 457
 |||||

Y 1424 tgaataataaagccatttgaanaacggaagaatlagatactaaatnactttracataatc 1403

[illegible]

Y 1544 ttctcgtgaatttcttcctcgtatatttaaacaatacgtagttacaaatat 1603
| | | | | | | | | | | | | | |
638 TCTATTTAATAAATAAGTAATACCTTTAATAGAATA-----ATAATTAAAT 689

1663 *ctcdaagaatcgttccaattggaagggcaagggccaggttcgacattttcaacctgcat* 1663
1664 *ttatataataataatcttcttgaagctattttccattttttttttttaaaggtat* 749

750 ATATCTTAAAAAGATTAAATTAAATTACATAAT 783

Sequence 541, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Rohmann, Paul

APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jürgen
APPLICANT: Knechtle, Philipp
APPLICANT: Reibischung, Corinne

TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSTYII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESSES:
ADDRESSEE: No. 6239264artis Corporation

CITY: Research Triangle Park

Tue Jul 2 10:13:51 2002

us-09-674-235-19.rni

Page 6

ADDRESS: THE WEBB LAW FIRM
 STREET: 700 KOPPERS BUILDING, 436 SEVENTH AVENUE
 CITY: PITTSBURGH
 STATE: PENNSYLVANIA
 COUNTRY: UNITED STATES OF AMERICA
 ZIP: 15219-1818
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" FLOPPY DISK
 COMPUTER: Midwest Micro 486-50
 OPERATING SYSTEM: DOS
 SOFTWARE: WORDPERFECT 6.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/451,405A
 FILING DATE: 26-MAY-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/965,273
 FILING DATE: 15-JAN-1993
 INFORMATION FOR SPO ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 731
 TYPE: NUCLEIC ACID
 STRANDEDNESS: SINGLE
 TOPOLOGY: UNKNOWN
 US-08-451-405A-2

Query Match	2.6%	Score	44.6;	DB	1;	Length	731
Best Local Similarity	43.7%;						
Matches	197;	Conservative	0.026;				

[illegible]

RESULT 10
US-07-867-106-2/c
; Sequence 2, Application US/07867106
; Patent No. 5389526
; GENERAL INFORMATION:
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular

```

? ? ? ? ? TITLE OF INVENTION: Slime Moulds of the Genus Dictyostellium
? ? ? ? ? NUMBER OF SEQUENCES: 19
? ? ? ? ? CORRESPONDENCE ADDRESS:
? ? ? ? ? ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526rIs
? ? ? ? ? STREET: One Liberty Place 46th Floor
? ? ? ? ? CITY: Philadelphia
? ? ? ? ? STATE: PA
? ? ? ? ? COUNTRY: USA
? ? ? ? ? ZIP: 19103
? ? ? ? ? COMPUTER READABLE FORM:
? ? ? ? ? MEDIUM TYPE: Floppy disk
? ? ? ? ? COMPUTER: IBM PC Compatible
? ? ? ? ? OPERATING SYSTEM: PC-DOS/MS-DOS
? ? ? ? ? SOFTWARE: PatentIn Release #1.0, Version #1.25
? ? ? ? ? CURRENT APPLICATION DATA:
? ? ? ? ? APPLICATION NUMBER: US/07/867,106
? ? ? ? ? FILING DATE: 19920625
? ? ? ? ? PRIOR APPLICATION DATA:
? ? ? ? ? APPLICATION NUMBER: AU PJ 7187
? ? ? ? ? FILING DATE: 02-NOV-1989
? ? ? ? ? ATTORNEY/AGENT INFORMATION:
? ? ? ? ? NAME: Feeoney, Joanne Longo
? ? ? ? ? REGISTRATION NUMBER: RICE-0002
? ? ? ? ? TELECOMMUNICATION DOCKET NUMBER: 35,134
? ? ? ? ? TELEPHONE: 215-568-3100
? ? ? ? ? TELEFAX: 215-568-3439
? ? ? ? ? INFORMATION FOR SEQ ID NO: 2:
? ? ? ? ? SEQUENCE CHARACTERISTICS:
? ? ? ? ? LENGTH: 5852 base pairs
? ? ? ? ? TYPE: NUCLEIC ACID
? ? ? ? ? STRANDEDNESS: single
? ? ? ? ? TOPOLOGY: linear
? ? ? ? ? MOLECULE TYPE: DNA (genomic)
? ? ? ? ? ANTI-SENSE: NO
? ? ? ? ? FEATURE:
? ? ? ? ? NAME/KEY: CDS
? ? ? ? ? LOCATION: 2378..5038
? ? ? ? ? FEATURE:
? ? ? ? ? NAME/KEY: CDS
? ? ? ? ? LOCATION: 2378..5038
? ? ? ? ? US-07-867-106-2

Query Match          2.6%   Score 44; DB 1; Length 5852;
Best Local Similarity 49.7%; Pred. No. 0.088;
Matches 196; Conservative 0; Mismatches 190; Indels 8; Gaps 3

QY 1192 ttctgggaagtcctttaaaagtctacaattacataataatctactgtagataagaatga 1251
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2378 TTTTAAATCTTGCAATTAAATAATTAAACCAAACCTATTTAATTAATAAATAAATAAGA 2319
QY 1252 ttcgaacaacaagaanaatcacagaatttagagtggtgcggtcgtgtgtaagaaccat 1311
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2318 AT-TTAAACAATTAATAAATAATTAATAAGAAAGTGGGAAAAATTAATAAATAATT 1260
QY 1312 gtgatgaattcataaaggctgcgaagaatcaaacaaactactgtcatgcaccagaatcaa 1371
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2259 AAAAATTAATAAATAAATAAAGAAATTTTAAAAAAAATAAATAAATAAATAAATAAATAA 2200
QY 1372 aataaatccagaatfagagactataataatgcatltaatacatgactacttgcacataa 1431
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2199 AATTAATATTAATTAATAATTCATGCCAAAACGTGAATAATTTGATATATTAATTCACA 2140
QY 1433 taagccattggaaaaaggaagaatfatagatactaataaactgaactactctcttgcnaat 1491
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2139 T-----TATATTAATTAAGGTAAATTAATGATGAGAAATTAATTTTAATTTATTAATA 2085
QY 1492 acagatcactaaa--tgatgttagttactttccaatggtgynaatttaactacttttctt 1549
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2084 AAATCAAAAAAAAAACCAGTAATTAATTAATTAATGAAGGTTTTTTTTTTTTTTTTTTT 2035

```


COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

Query	March	2.68;	Score	43.6;	DB	1;	Length	1511;
Best Local Similarity	45.5%;	Pred.	No.	0.063;				
Matches	282;	Conservative	0;	Mismatches	324;	Indels	14;	Gaps
QY	1076	caaccagcaagaacatatactatagatcacatacaatcccaattataattatgycattaattg	1135					
Db	825	CAATGAATAAAAAATAAAATATCAAAAATGGATTCTAAATTCATATTAATTTAATA	884					
QY	1136	catttattccaaataactctgtgggacatglaactctcttgagaaactgatactttt	1195					
Db	885	AAATATTTTAAATAATTAATTAATTTAAAAAAAATAATATATAACAGACATATATGTTATAT	944					
QY	1196	gggaagtccttataaaagttacaaatttcataataactagtagaagaatgattca	1255					
Db	945	ATTAAATATTTAAAAAAATAGTGAATTTAGAGAATTCATATATATATATGATATAT	1004					
QY	1256	gaacacaaagaacatcacgaattggaatgctgcgcgcggtgctatgaagccaactgta	1315					
Db	1005	AATATTTTAAATAATATCCACAAAATATTTAAAGTTTATATATTTCA-----AT	1055					
QY	1316	tgaattcataaagttgcaaaagtcacaaacaactactgcatcagcaaccaagaatcaaat	1375					
Db	1056	TTAAATATTTATTAATTTTAAATTTTATACAAAATTTAAAAATATATACATATTTGAAATA	1115					
QY	1376	aatccagaataagagactatataatgcacttatacatagatcatttgacataatgaag	1435					
Db	1116	TCTTATTAACAAAATAATGCAATATAGTAATTTACTACACATCTGTATAGAAATTTT	1175					
QY	1436	ccattggaacacggaaagatfagataactaataaacttgactctcttgytaataacag	1495					
Db	1176	AATTGTGAATCATTAATATATAAGACTATA-----ATTATTTAAATAATTTAGTAATTT	1231					
QY	1496	tcactaataatgatcttagtctacttccatiggtggaattcaattactcttcttgtaatt	1555					
Db	1232	AAAAAAATTAATATATCTAAAAATAAATTTGGTACTTAAATAAGTTTTCCTATTTAG	1291					

FEATURE:	CDS
NAME/KEY:	

APPLICANT: Reibischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF *NEBRYA COCCINATA*

TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSEYII
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSEYII
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSEYII

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: July 1, 2002, 15:45:51 Search time 316.64 seconds
(without alignments) 9245.007 Million cell updates/sec

Title: US-09-674-235-19
Contact record: 1705

Scoring table: IDENTITY_NUC
Gapop 10.0 / Gapext 1.0

```
Searched:      1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

- 1: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1981.DAT.*
- 2: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1982.DAT.*
- 3: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1983.DAT.*
- 4: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1984.DAT.*
- 5: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1985.DAT.*
- 6: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1986.DAT.*
- 7: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1987.DAT.*
- 8: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1988.DAT.*
- 9: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1989.DAT.*
- 10: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1990.DAT.*
- 11: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1991.DAT.*
- 12: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1992.DAT.*
- 13: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1993.DAT.*
- 14: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1994.DAT.*
- 15: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1995.DAT.*
- 16: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1997.DAT.*
- 17: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1997.DAT.*
- 18: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1998.DAT.*
- 19: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA1999.DAT.*
- 20: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA2000.DAT.*
- 21: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA2001A.DAT
- 22: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA2001B.DAT
- 23: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA2002.DAT.*
- 24: /SIDS1/gcgata/hold-geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1705	100.0	1705	21	AAZ38317	Human transmembran
2	1000	58.7	1030	22	AAH98847	Human EST-derived
3	804	47.2	804	21	AAZ38316	Human transmembran
4	801.8	47.0	2684	22	AAK85824	HOAT5 DNA. Homo s
5	773	45.3	783	22	AAK70136	Human immune/Haema
6	761	44.6	782	22	AAK70138	Human immune/Haema
7	759.4	44.5	782	22	AAK70137	Human immune/Haema
8	738.6	43.3	1316	22	AAZ55043	Nicotinamide sequen
9	687	40.3	1355	22	ABA08976	Human secreted p

10	399	23.4	1375	22	AA1599215	Human poly(ADP-ribose)
11	399	23.4	1986	22	AA059631	Human transporter
12	395.8	23.2	2027	21	AA6618521	CDNA encoding a hu
13	378	22.2	1977	22	AA685823	hOAT4 DNA. Homo s
14	232.8	13.7	1132	22	AA161001	Human poly(ADP-ribo
15	218.4	12.8	2865	22	AA012566	Human transporter
16	208.8	12.2	2455	22	AA017476	Human transporter
17	180.8	10.6	2047	22	AA017480	Human protein havi
18	180.8	10.6	2194	22	AA012607	Human organic anio
19	180.8	10.6	2210	22	AA032614	Human foetal liver
20	161.8	9.5	540	22	ABA63164	Probe #8860 for ge
21	161.8	9.5	540	22	ABA63034	Human foetal liver
22	161.8	9.5	540	22	AAK11579	Human brain expres
23	161.8	9.5	540	22	AAK37346	Human bone marrow
24	161.8	9.5	540	22	AA118114	Probe #8107 for ge
25	161.8	9.5	540	22	AA143191	Probe #11877 used
26	145.2	8.5	447	22	AA107074	Human reproductive
27	145.2	8.5	9370	22	AA107075	Human foetal liver
28	136.2	8.0	339	22	ABAV5594	Probe #18672 for g
29	136.2	8.0	339	22	ABAA4026	Human brain expres
30	136.2	8.0	339	22	ABAA4178	Human bone marrow
31	136.2	8.0	339	22	AAK50227	Probe #17241 for g
32	136.2	8.0	339	22	AAK50227	Probe #24865 used
33	136.2	8.0	339	22	AA127308	Rat EST-derived co
34	134.4	7.9	685	22	AAH56179	Human CDNA 5'-end
35	108	6.3	789	22	AAK92269	Human CDNA clone r
36	108	6.3	789	22	AAK93561	Human full-length
37	108	6.3	1871	22	AAK94561	Human full-length
38	98.8	5.8	280	21	AAAI0196	Rat liver toxicolo
39	98.2	5.8	419	22	AAI02433	Human reproductive
40	88.6	5.2	2294	20	AAV79584	Rat organic anion
41	85.2	5.0	379	22	AA184356	Human polynucleoti
42	83.8	4.9	2123	21	AAZ29300	Human organic acid
43	83.8	4.9	2127	22	AAAC8519	HOAT1 DNA. Homo s
44	83.8	4.9	2171	20	AAV79585	Human organic acid
45	81	4.8	1638	22	AAAC8044	Mouse organic acid

ALIGNMENTS

RESULT	1
ID	AA238317 standard; cDNA; 1705 BP.
XX	AA238317;
AC	
XX	
DT	09-FEB-2000 (first entry)
XX	
DE	Human transmembrane protein cDNA clone HP02000.
XX	
XX	HP02000; transmembrane domain; liver; expression; homology;
XX	organic cation transporter; drug excretion; antibody; assay reagent;
KW	diagnostic marker; primer; probe; antisense; gene therapy;
KW	agonist; antagonist; ligand; therapeutic; ds.
XX	
OS	Homo sapiens.
XX	
XX	
FH	Key
FT	Location/Qualifiers
FT	187..993
FT	/*tag= a
FT	/*product= "Human transmembrane protein HP02000"
XX	
PN	W09955862-A2.
XX	
PD	04-NOV-1999.
XX	
PF	27-APR-1999; 99WO-JP02226.
XX	
PR	28-APR-1998; 98JP-0119395.
XX	
PA	(SAGA) SAGAMI CHEM RES CENT.
PA	(PROT-) PROTEGENE INC.

XX Kato S, Kimura T;
 PI
 DR WPI: 2000-023358/02.
 DR P-PSDB: AAY52386.
 XX

Human proteins with transmembrane domains, involved in control of cell proliferation and differentiation, useful for treating e.g. cancer or inflammation

Claim 4; Page 89-91; 114pp; English.

This sequence represents the human cDNA clone HP02000 which encodes a 32 kD protein with two putative transmembrane domains. The cDNA was isolated from a human liver cell cDNA library, and from tissue localisation studies has been found to be expressed only in the liver. The protein has homology with the rat organic cation transporter (EMBL Accession No. Y09945) which is involved in drug excretion, and may have a similar function. The protein may be used to raise specific antibodies, as assay reagents, as diagnostic tissue markers, for the isolation of cognate receptors, ligands and binding proteins, and as biologically active agents. Nucleotides encoding the protein may be used as primers and probes or antisense molecules, and in gene therapy. Cells transformed with these nucleotides may be used to screen for agonists and antagonists which are potentially useful therapeutically.

Sequence 1705 BP; 559 A; 305 C; 344 G; 497 T; 0 other;

Query Match 100.0%; Score 1705; DB 21; Length 1705;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 aagaactggaagagctcttcctacacgctgatactgacgtgagtcgcggccatg 60
DB 1 aagaactggaagagctcttcctacacgctgatactgacgtgagtcgcggccatg 60
QY 61 atgagaatgatgcgagagatcaatacaagcttaattgataataaagaaatttt 120
DB 61 atgagaatgatgcgagagatcaatacaagcttaattgataataaagaaatttt 120
QY 121 tctcccttggaattatcttcctgaagaagcattgctcctctctggggcagctgtca 180
DB 121 tctcccttggaattatcttcctgaagaagcattgctcctctctggggcagctgtca 180
QY 181 caatcaatgagccttgagagctctgagtcgaagtgaggagccttgaggattcagatg 240
DB 181 caatcaatgagccttgagagctctgagtcgaagtgaggagccttgaggattcagatg 240
QY 241 ctctatcgtgttattatcttcctctcctcatgtaataatccatatactgctagaag 300
DB 241 ctctatcgtgttattatcttcctctcctcatgtaataatccatatactgctagaag 300
QY 301 aacttgctgcagcattccctgctgcatcgttgcgtgctccatgctggaataataact 360
DB 301 aacttgctgcagcattccctgctgcatcgttgcgtgctccatgctggaataataact 360
QY 361 ggatctggttaataactggaatccctcagtggaagtgcaccttgaagatcctatcca 420
DB 361 ggatctggttaataactggaatccctcagtggaagtgcaccttgaagatcctatcca 420
QY 421 ctgaactaaatctgagagccagagaagtgcgtcgttgccttgcaccccaatgagactt 480
DB 421 ctgaactaaatctgagagccagagaagtgcgtcgttgccttgcaccccaatgagactt 480
QY 481 cttaactgaatgagactatccacagacaagtgaagagacacagaacctgtgtgat 540
DB 481 cttaactgaatgagactatccacagacaagtgaagagacacagaacctgtgtgat 540
QY 541 ggcctggatataatgatacaaaactacttccctcgacacattgagataatgagactgta 600
DB 541 ggcctggatataatgatacaaaactacttccctcgacacattgagataatgagactgta 600

```

```

QY 601 tctgatatataatgaactgaatcagtggttcaatctcctactctgactggaatgctgtg 660
DB 601 tctgatatataatgaactgaatcagtggttcaatctcctactctgactggaatgctgtg 660
QY 661 ggaagcatatagtggtgcaatcctcagacagagtgctggtggaatcctgctggtg 720
DB 661 ggaagcatatagtggtgcaatcctcagacagagtgctggtggaatcctgctggtg 720
QY 721 atataccaataaactatagatgaggtctaaagcacttagaagattgcacacacaaat 780
DB 721 atataccaataaactatagatgaggtctaaagcacttagaagattgcacacacaaat 780
QY 781 ggaataaagaatgctgaaagaacccctgaacatagaggtgtgaagatccacatgcagag 840
DB 781 ggaataaagaatgctgaaagaacccctgaacatagaggtgtgaagatccacatgcagag 840
QY 841 ggaactgagatgcagacagacacaaactactgctgtgactgttccgcaacccagatg 900
DB 841 ggaactgagatgcagacagacacaaactactgctgtgactgttccgcaacccagatg 900
QY 901 cgtaaagagatcgtatctggttattttgagaataaaatctcaaggaaagacataaa 960
DB 901 cgtaaagagatcgtatctggttattttgagaataaaatctcaaggaaagacataaa 960
QY 961 aatgattctacacaaagtgacacaaatttaagaagccttcataagctgattggtggg 1020
DB 961 aatgattctacacaaagtgacacaaatttaagaagccttcataagctgattggtggg 1020
QY 1021 aatctcagaataaataatagcagaaagaaacacacacagaggtttttccctacaac 1080
DB 1021 aatctcagaataaataatagcagaaagaaacacacacagaggtttttccctacaac 1080
QY 1081 agcaagaacatataatagatcacgaatcctaataatattatgacattatgttcatt 1140
DB 1081 agcaagaacatataatagatcacgaatcctaataatattatgacattatgttcatt 1140
QY 1141 taattcaaatataatctgtgtggtgacatgtaattctctgagcaatctgataatttggaa 1200
DB 1141 taattcaaatataatctgtgtggtgacatgtaattctctgagcaatctgataatttggaa 1200
QY 1201 gtcccttaaaagttacaatattacaataaattactagtagataagaatgattcagaac 1260
DB 1201 gtcccttaaaagttacaatattacaataaattactagtagataagaatgattcagaac 1260
QY 1261 aaaaagaataacacagaaattgagatgctgctgctgtgtatgaagacacatgcatgaat 1320
DB 1261 aaaaagaataacacagaaattgagatgctgctgctgtgtatgaagacacatgcatgaat 1320
QY 1321 tcaataagttgcaaaagtcacaaacatctgtaacatgcaacacagaaatcaaaataatcc 1380
DB 1321 tcaataagttgcaaaagtcacaaacatctgtaacatgcaacacagaaatcaaaataatcc 1380
QY 1381 agaaataagagactataataatgacatttaatacatgatacttttgcacataaagcatt 1440
DB 1381 agaaataagagactataataatgacatttaatacatgatacttttgcacataaagcatt 1440
QY 1441 ggaataacgaaagatagataactataaataactgatactcttctgttaatacagctact 1500
DB 1441 ggaataacgaaagatagataactataaataactgatactcttctgttaatacagctact 1500
QY 1501 aaatgagtgtaattcttccatgctggaatcttaataactcttctgtgaatttttc 1560
DB 1501 aaatgagtgtaattcttccatgctggaatcttaataactcttctgtgaatttttc 1560
QY 1561 tctcgtatattttaaataaataatagctgtgtatagttacataatataaagatatgttca 1620
DB 1561 tctcgtatattttaaataaataatagctgtgtatagttacataatataaagatatgttca 1620
QY 1621 aattgaagggcaagggccaggttcagcaatttcaaaactgtatgataatataaataa 1680
DB 1621 aattgaagggcaagggccaggttcagcaatttcaaaactgtatgataatataaataa 1680

```

OY 1681 actataaataaataatatttc 1705
 |||
 DB 1681 actataaataaataatatttc 1705

RESULT 2
 AAH98847
 ID AAH98847 standard; cDNA; 1030 BP.

AAH98847;
 12-OCT-2001 (first entry)

Human EST-derived coding sequence SEQ ID NO: 704.

Human: sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 diagnostics; forensic test; gene mapping; genetic disorder;
 biodiversity; gene therapy; nutrition; ss.

Homo sapiens.
 W0200154477-A2.

02-AUG-2001.

25-JAN-2001; 2001WO-US02687.

25-JAN-2000; 2000US-0491404.

17-JUL-2000; 2000US-0617746.

03-AUG-2000; 2000US-0631451.

15-SEP-2000; 2000US-0663870.

(HYSEQ) HYSEQ INC.

Tang XT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 Cao Y, Dermanac RA, Zhang J, Werhman T;

WPI: 2001-476164/51.

P-PSDB; AAM24188.

Isolated polypeptide for treatment of diseases, diagnostics, raising
 antibodies and research use -

Claim 1; Page 638; 1275pp; English.

The present invention provides the protein and coding sequences of novel
 proteins from a variety of organisms, including human, dog, cat, horse,
 cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 urchin and tomato. These were derived from expressed sequence tags (ESTs)
 from the organism of interest. They can be used in diagnostics,
 forensics, gene mapping, identification of mutations, to assess
 biodiversity and for nutritional purposes. The present sequence is a cDNA
 of the invention.

Sequence 1030 BP; 289 A; 217 C; 247 G; 277 T; 0 other:

Query Match 58.7%; Score 1000; DB 22; Length 1030;
 Best Local Similarity 99.8%; Pred. No. 1.7e-218;
 Matches 1022; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

OY 8 gaggaagctcttcactacagcgtatgacgtgtgagtcggcgccatgagatgaa 67
 |||
 DB 7 gaggaagctcttcactacagcgtatgacgtgtgagtcggcgccatgagatgaa 66
 |||
 OY 68 attatgagggatcaataaagcttaattgaaataaagaataatttcctcct 127
 |||
 DB 67 attatgagggatcaataaagcttaattgaaataaagaataatttcctcct 126
 |||
 OY 128 ttgaactatctccgtaagcattgtgctcctctggtgggtgacgtgttcacatcaa 187
 |||
 DB 127 ttgaactatctccgtaagcattgtgctcctctggtgggtgacgtgttcacatcaa 186
 |||

OY 188 tggccttgaggagctctctgagtcagcttgaggccttgaggatctcaatgctcacc 247
 |||
 DB 187 tggccttgaggagctctctgagtcagcttgaggccttgaggatctcaatgctcacc 246
 |||
 OY 248 tggctttatctctcctctctcaatgtatattatccctcatatctgtagaacttg 307
 |||
 DB 247 tggctttatctctcctctctcaatgtatattatccctcatatctgtagaacttg 306
 |||
 OY 308 ctgcagcattccctggtcatctgtctggtgtccacatgctgagacaataactgactg 367
 |||
 DB 307 ctgcagcattccctggtcatctgtctggtgtccacatgctgagacaataactgactg 366
 |||
 OY 368 gtaatgaaactggaactcctcagtgagaatgacctcttgagaatctctaccactagact 427
 |||
 DB 367 gtaatgaaactggaactcctcagtgagaatgacctcttgagaatctctaccactagact 426
 |||
 OY 428 caaatctgagggcagagaagctgctgctctgtccatccctcagtgagactctccacc 487
 |||
 DB 427 caaatctgagggcagagaagctgctgctctgtccatccctcagtgagactctccacc 486
 |||
 OY 488 tgaatgagactatccacagcacaagtgagagcagacacagaccctggtgagctgctgg 547
 |||
 DB 487 tgaatgagactatccacagcacaagtgagagcagacacagaccctggtgagctgctgg 546
 |||
 OY 548 tataatgacaagcttactctccttcgaccattgtgactaagtggagacctgtagatgatt 607
 |||
 DB 547 tataatgacaagcttactctccttcgaccattgtgactaagtggagacctgtagatgatt 606
 |||
 OY 608 atcagtcactgaaatcagtgcttcaattcctactcttgaactggaatgctgtgtgagagca 667
 |||
 DB 607 atcagtcactgaaatcagtgcttcaattcctactcttgaactggaatgctgtgtgagagca 666
 |||
 OY 668 tcaatggtgagcattgtctcagagagctggtgtggaatctgctgctgtgtgataatca 727
 |||
 DB 667 tcaatggtgagcattgtctcagagagctggtgtggaatctgctgctgtgtgataatca 726
 |||
 OY 728 ccataaactagatgagagcttaagagcacttagaagaagtgcagacacaaatgagataa 787
 |||
 DB 727 ccataaactagatgagagcttaagagcacttagaagaagtgcagacacaaatgagataa 786
 |||
 OY 788 agaatgctg-aagaaaccccggaacatagaggtgtgtaagatccacacatgcagagagctg 846
 |||
 DB 787 agaatgctg-aagaaaccccggaacatagaggtgtgtaagatccacacatgcagagagctg 846
 |||
 OY 847 gatcagacagacaaactactctgtgtgactgtgttcgcgcaacccagatgagctgaa 906
 |||
 DB 847 gatcagacagacaaactactctgtgtgactgtgttcgcgcaacccagatgagctgaa 906
 |||
 OY 907 aggatctgatactctgtatctttgag-aaaaaatctaaaggaagaagcataaaatga 965
 |||
 DB 907 aggatctgatactctgtatctttgag-aaaaaatctaaaggaagaagcataaaatga 966
 |||
 OY 966 ttgctacacaaagtgacaaattttaagaagccttcataagctgattgtgtgggaatt 1025
 |||
 DB 967 ttgctacacaaagtgacaaattttaagaagccttcataagctgattgtgtgggaatt 1026
 |||
 OY 1026 caga 1029
 |||
 DB 1027 caga 1030

RESULT 3
 AAZ38316
 ID AAZ38316 standard; cDNA; 804 BP.
 AAZ38316;
 09-FEB-2000 (first entry)
 Human transmembrane protein cDNA clone HP02000 coding sequence.
 HP02000; transmembrane domain; liver; expression; homology;

27 tcaaatctgaagccagaagaatattcaccatttatgacac

Sun; W;

R
WPI; 2001-367057/38.

PR	31-JAN-2000;	2000US-0179065
PR	04-FEB-2000;	2000US-0180628
PR	24-FEB-2000;	2000US-0184654
PR	02-MAR-2000;	2000US-0185650
PR	16-MAR-2000;	2000US-0189674
PR	17-MAR-2000;	2000US-0190076
PR	18-MAR-2000;	2000US-0198123
PR	19-MAY-2000;	2000US-0205515
PR	07-JUN-2000;	2000US-0209467
PR	28-JUN-2000;	2000US-0214686
PR	30-JUN-2000;	2000US-0215345
PR	07-JUL-2000;	2000US-0216647
PR	07-JUL-2000;	2000US-0216880
PR	11-JUL-2000;	2000US-0217487
PR	14-JUL-2000;	2000US-0217496
PR	26-JUL-2000;	2000US-0218290
PR	26-JUL-2000;	2000US-0220963
PR	14-AUG-2000;	2000US-0220964
PR	14-AUG-2000;	2000US-0224518
PR	14-AUG-2000;	2000US-0224519
PR	14-AUG-2000;	2000US-0225213
PR	14-AUG-2000;	2000US-0225214
PR	14-AUG-2000;	2000US-0225266
PR	14-AUG-2000;	2000US-0225267
PR	14-AUG-2000;	2000US-0225267
PR	14-AUG-2000;	2000US-0225268
PR	14-AUG-2000;	2000US-0225270
PR	14-AUG-2000;	2000US-0225447
PR	14-AUG-2000;	2000US-0225457
PR	14-AUG-2000;	2000US-0225758
PR	18-AUG-2000;	2000US-0225759
PR	22-AUG-2000;	2000US-0226279
PR	22-AUG-2000;	2000US-0226661
PR	22-AUG-2000;	2000US-0226668
PR	22-AUG-2000;	2000US-0227182
PR	30-AUG-2000;	2000US-0227009
PR	01-SEP-2000;	2000US-0228294
PR	01-SEP-2000;	2000US-0229287
PR	01-SEP-2000;	2000US-0229343
PR	01-SEP-2000;	2000US-0229344
PR	05-SEP-2000;	2000US-0229345
PR	05-SEP-2000;	2000US-0229509
PR	06-SEP-2000;	2000US-0229517
PR	06-SEP-2000;	2000US-0230437
PR	08-SEP-2000;	2000US-0230438
PR	08-SEP-2000;	2000US-0231142
PR	08-SEP-2000;	2000US-0231143
PR	08-SEP-2000;	2000US-0231144
PR	08-SEP-2000;	2000US-0231145
PR	08-SEP-2000;	2000US-0231146
PR	08-SEP-2000;	2000US-0231147
PR	08-SEP-2000;	2000US-0231148
PR	12-SEP-2000;	2000US-0232081
PR	14-SEP-2000;	2000US-0231968
PR	14-SEP-2000;	2000US-0232297
PR	14-SEP-2000;	2000US-0232398
PR	14-SEP-2000;	2000US-0232399
PR	14-SEP-2000;	2000US-0232400
PR	14-SEP-2000;	2000US-0232401
PR	14-SEP-2000;	2000US-0233063
PR	21-SEP-2000;	2000US-0233065
PR	21-SEP-2000;	2000US-0234223
PR	21-SEP-2000;	2000US-0234273
PR	25-SEP-2000;	2000US-0234274
PR	25-SEP-2000;	2000US-0234997
PR	26-SEP-2000;	2000US-0234998
PR	27-SEP-2000;	2000US-0234834
PR	27-SEP-2000;	2000US-0235836
PR	29-SEP-2000;	2000US-0236337
PR	29-SEP-2000;	2000US-0236367
PR	29-SEP-2000;	2000US-0236368
PR	29-SEP-2000;	2000US-0236369
PR	29-SEP-2000;	2000US-0236370
PR	02-OCT-2000;	2000US-0236802

XX	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	02-OCT-2000;	2000US-0237040.
PR	02-OCT-2000;	2000US-0239935.
PR	13-OCT-2000;	2000US-0239937.
PR	13-OCT-2000;	2000US-0240960.
PR	20-OCT-2000;	2000US-0241221.
PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241786.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	20-OCT-2000;	2000US-0244617.
PR	01-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246509.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	08-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	17-NOV-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	01-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251899.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
XX		
PI	Rosen CA, Barash SC, Ruben SM;	
XX		
XX	WPI; 2001-483426/52.	
DR	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating cancers and	
PT	metastasis -	

	xx	Disclousre; SEQ ID NO 24948; 3071pp + Sequence listing; English.
	xx	
	cc	AAK64702 encode the human immune/hematopoietic antigen (I)
	cc	amino acid sequences given in AAK82170 to AAM91921. (I) have cytosolic
	cc	activity, and can be used in gene therapy and vaccine production. (I)
	cc	proteins and polynucleotides may be used in the prevention, diagnosis and
	cc	treatment of diseases associated with inappropriate (I) expression. For
	cc	example, they may be used to treat disorders associated with decreased
	cc	expression by rectifying mutations or deletions in a patient's genome
	cc	that affect the activity of (I) by expressing inactive proteins or to
	cc	supplement the patients own production of (I). Additionally, (I)
	cc	polynucleotides may be used to produce the secreted (I), by inserting the
	cc	protein. (I) proteins and polynucleotides may be used to prevent,
	cc	diagnose and treat immune/hematopoietic-related diseases, especially
	cc	cancers and cancer metastases of hematopoietic-derived cells. AAK64703
	cc	to AAK87694 represent human Immune/hematopoietic antigen genomic
	cc	sequences from the present invention. AAK54942 to AAK54950 and AAM82169
	cc	represent sequences used in the exemplification of the present invention.
	xx	
SQ	Sequence 783 BP; 242 A; 119 C; 105 G; 317 T; 0 other;	
	Query Match	45.3%; Score 773; DB 22; Length 783;
	Best Local Similarity	100.0%; Pred. No. 9.1e-167;
	Matches 773; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Db	933 aaaaaaatcgaagaaagcataaaatgatgtgctacaaaagtgcacaatttta	992
Db	776 AAAAAAAAAATCCAGAAAAGGCGATAAATGATTGTCTACCAAAAGTGACCAAATTTN	717
Oy	agaagccttcatacgctcgattgtgtggggaattcagaaaaaaaaaatacagaanaagaca	1052
Db	716 AGAAGCCTTCATGAGCTGATGGTGGGAATTCAGAAAAAAATACAGAAAAAGMACA	657
Oy	caccagaaggatttttccctacaaccaagacaacatatattaagtaataigaatcca	1112
Db	656 CACCGAAGAGGTTTTTTCCTACAACGACGAGAACAATAATTAGTAACATGAAATCTCA	597
Oy	attataattatggcatatttgcaatttatctcaaaatlaacttgtgtggagacatgtaat	1172
Db	596 ATTATAAATTATGGCATTAATATTGCAATTTATTTCAAAATTAACCTGTGGGACATGTAAAT	537
Oy	ctcttgaagcaatctgatalttlttgggaagtcoclttaaaagtltaacaattatataana	1232
Db	536 CTCTTGAGCAATCTGATATTTTGGGAAGTCCTTTAAAAAGTTACAAATTTATCAATAAA	477
Oy	ttacttagtagataagatgattcagaacaacaaagaataacacagaattggatgtgctgtg	1292
Db	476 TTACTAGTAGATTAAGATGATTCAGAAACAAAAGAAATCACGAATTTAGCATGTGCTGG	417
Oy	ctgtgtatgaagcacacatgfatgtaattcataaaglttgcaaaagltcaaacatactgt	1352
Db	416 CTGTGTATGAGCAACCATGTGATGAATTCATTAAGATTGCCAAAAGTCAAAACAAATACGT	357
Oy	acatgcaaccagaataccaataataatccagaatagagagccatataatgatcatiaata	1412
Db	356 ACATGCACACGAAATCAAAATTAATCCAGAAATTAAGACCTRTATTAATTCGATTTAAAT	297
Oy	catgatacttttgaacataaagccatitggnaaacggaagatltagatactaataaocat	1472
Db	296 CATGATACCTTTGACATTAATTAAGCCATTTGGAANAACGGAAGATTAGACTAAATAACAT	237
Oy	tgacatactctttgtaaatcacgtacctaaatgaaatgattagttacttttcgatgtggaat	1532
Db	236 TGACATACCTTTGTAATTAATACAGTCACATAATGATGTATTAATCTTTTCATGTGGAAAT	177
Oy	tttaactacttttcttctgtaattttctctctgtaatttttaaacaacatagctgtlata	1592
Db	176 TTTAATTAATCTTTCTTTGTAATTTTCTCTCTGTATTAATTTTAAACAAATAGCTGTATA	117
Oy	gttacaatatataagaatatgttgcnaattgaagaggcaagggccaggttccagcaattt	1652

|||||
Db 116 GTTAAATATATTAAGATATGTCATATTAAAGGCGAAGCCAGGTCAGCAATTT 57
Qy 1653 tcaactgtatgtacacttaataaactaataaataataataatatttc 1705
Db 56 TCAAACTGATGTACATTTAAATAAATACTAATAATTAATAATATATTTTC 4
RESULT 6
AAK70138/C
ID AAK70138 standard; DNA; 782 BP.
XX
AC AAK70138;
XX
DT 06-NOV-2001 (first entry)
XX
Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24950.
XX
Human; immune: haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
Homo sapiens.
XX
PN W0200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0190876.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-020515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225277.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226679.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231245.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234498.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.

PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249289.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PR
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 DR WPI: 2001-483426/52.
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure: SEQ ID NO 24950; 3071bp + Sequence Listing: English.

CC AAK4951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytosolic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 CC
 SQ Sequence 782 BP; 241 A; 119 C; 105 G; 317 T; 0 other;

Query Match 44.6%; Score 761; DB 22; Length 782;
 Best Local Similarity 99.9%; Pred. No. 5e-164; 0; Indels 1; Gaps 1;
 Matches 772; Conservative 0; Mismatches

QY 933 aaaaaaatctcaagaaagcctaaataatgattgctacacaaagtgcacaaatttla 932
 DB 775 AAAAAAAATCTCAAGGAAAAAGCCATTAATGATTCCTACCAAAAGACCAATTTTA 716
 QY 993 agaagcctcatgagctgattggtgggaaatcagaataaaatcagagaagaagaca 1052
 DB 715 AGAAGCCTCATGAGCTGATGTGTGGGCAAAATTCAGAAAAAAATACAGAAAAAGACA 656
 QY 1053 caccagaaggggttttttccctcaaacccagcaagacatatatagacatgaattcca 1112
 DB 655 CACCAGAAAGGGTTTTCCTTCAACACGCAAGAAACATATATGATACATGAATTCACA 596
 QY 1113 attataatgagcatatattgcatlittattcacaataactggtggagcatgtaatt 1172
 DB 595 ATTATATATGAGCATTAATTTGCAATTTATTTCAAAATTAACCTGTGGGACATGTAAT 536

QY 1173 ccttgagcaatctgatatattttgggaagtcctttaaaagttaacaattatcaataaa 1232
 DB 535 CTTTGAGCAATCTGATATTTTGGGAGTCTTTAAAGATTACAAATTTATCAATAAAA 476
 QY 1233 ttactagataagaatgatcagaacaagaataacagaagaattagagtgctg 1292
 DB 475 TTACTAGTATGATAGATGATTCAGAAACAAAAGAAATACAGAAATTAAGATGCGCTGG 416
 QY 1293 ctgggtatgaagcccatgagatgaatcataaagtgtcagaagtcacaataactgt 1352
 DB 415 CTGGGTATGAGCCACCATGTCATGATTAATTCATTAAGTTGCAAAAGTCAAAATACTGT 356
 QY 1353 acatgcaacagaatcaaaataaatcaagaataatgagagaccataataatgattaa 1412
 DB 355 ACATGCAACCCAGAAATCAAAATTAATTCAGAAATGAGACCTATATTAATGATTAATA 296
 QY 1413 catgatactttgacataaagccattggaagcgaagaattagataactaataaat 1472
 DB 295 CATGATCTTTGACATTAATTAAGCCANTGGAACGAAAGATTAGATCTAATTAACAT 236
 QY 1473 tgactatccttggtaaatcaagtcactaaatgattgattgtaactttccatggtgaat 1532
 DB 235 TGACTATCTCTTGTAAATACAGTCACTAATATGTTAGTTACTTTTCANTGTGGAAT 176
 QY 1533 tttaactcttttcttctgtaatttctcctgltatatttaacaatagctgata 1592
 DB 175 TTTAATTAATCTTTCTTTGTAAT-TTTCCTCTGTGATATTTTAACAAATGCTGATTA 117
 QY 1593 gttaacaatataaagatatcttcaaatggaagcgaagccagttcagcaatt 1652
 DB 116 GTTTCAATATATTAATTAATGATTTGTTCAATTAAGGCGCAAGCCAGGTTCAAGCAATT 57
 QY 1653 tcaaatgattatcatatttaataataactaataataataattatattc 1705
 DB 56 TCAAACTGATGATCATTTTAATTAATTAATTAATTAATTAATTAATTAATTTTC 4
 RESULT 7
 ID AAK70137 standard; DNA; 782 BP.
 XX AAK70137;
 AC
 XX 06-NOV-2001 (first entry)
 DT
 XX Human immune/hematopoietic antigen genomic sequence SEQ ID NO:24949.
 DE Human: immune; haematopoietic; immune/hematopoietic antigen; cancer;
 KW cytosolic; gene therapy; vaccine; metastasis; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01354.
 XX
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.

14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0220968.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0225759.
 PR 22-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 22-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0227009.
 PR 01-SEP-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 14-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249297.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 05-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-483426/52.
 Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 useful for preventing, diagnosing and/or treating cancers and
 metastasis -
 PS Disclosure; SEQ ID NO 24949; 3071pp + Sequence Listing; English.
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytosolic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I) by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,

CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK7694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK5442 to AAK4950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 CC
 XX Sequence 782 BP; 242 A; 119 C; 105 G; 316 T; 0 other;

Query Match 44.5%; Score 759.4; DB 22; Length 782;
 Best Local Similarity 99.7%; Pred. No. 1.1e-163;
 Matches 771; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

QY 933 aaaaaaacctcaagaaagcacaataatgctacacaaagtgcacaatttta 992
   |||||||
Db 775 AAAAAAACTCAAGAGAAAGGCATTAATATGCTCTACACAAAAGTGCATTTTA 716
   |||||||
QY 993 agaagcctcatgagctgtagtggggaattcagaaaaaaatacaggaagaaca 1052
   |||||||
Db 715 AGAAGCCTTCATGAGCTGATGTTGGGGAAATTCAGAAAAAAATACAGAAAGACA 656
   |||||||
QY 1053 caccagaagggttttttccctacaccagcaagaacatatatagatataatctca 1112
   |||||||
Db 655 CACCAAGAGGCTTTTTCCTCAACACAGCAAGACATATATTAGATGATGATCTCA 596
   |||||||
QY 1113 attataattatgcatatttgcattttatctcaaaataacttgggggaacatgta 1172
   |||||||
Db 595 ATTATATTATTGGCATTTATTTGCAATTTTCAAAATTAACCTTGGGGACATGTAAT 536
   |||||||
QY 1173 ctcttgagcaatctgatatcttgggaagctccttaaaagttaacaattatacaaa 1232
   |||||||
Db 535 CTCCTGAGCAATCTGATATTTTGGGAGGCTCTTAAAGTTTCAAAATTTATCAATAA 476
   |||||||
QY 1233 ttactagtaataagatgattcagaacaaagaaatacaagaattagagatgctg 1292
   |||||||
Db 475 TTACTATGATATGATGATTCAGAAACAAAGAAATTCACAGATTAAGATGTCGCTG 416
   |||||||
QY 1293 ctggtgtaagaaagcaccatgtagtaattcataaagtgcacaaagtcacaaactactgt 1352
   |||||||
Db 415 CTGCTGATGATGAGCAACATGATGATTAATTAAGTTGCAAAAGTCAAAACATATCTGT 356
   |||||||
QY 1353 acatgcaacagaatacaaaataatccagaatagagcctatataatgattataa 1412
   |||||||
Db 355 ACAAGCAACCGAAGATCAAAATTAATCCAGAAATAGACCTATATTAATGCAATTTAATA 296
   |||||||
QY 1413 catgaatactttgacataaagcattggaacacggaagaattagatactaaatacat 1472
   |||||||
Db 295 CATGATACCTTTTGACATTAATAGCCATTGGAACGGAAGTTATGATTAATAATACAT 236
   |||||||
QY 1473 tgactatcctctgtaaatagcagtaactaaatgattgattactttccatggtggaat 1532
   |||||||
Db 235 TGACTATCTCTTTTAATACAGTCACATAAATGATGTTCTTTCCATGAGTGGAAT 176
   |||||||
QY 1533 tttaatacttttcttgtaattttctcctgtaatttaacaataatgctgata 1592
   |||||||
Db 175 TTTAATTACTTTTCTTTGTAATTTTCTCTCTGATATTTTAACAATAATAGCTGTAATA 117
   |||||||
QY 1593 gttacaatatataaagatatgttcaaatgtgaagggcaagccaggttcagcaattt 1652
   |||||||
Db 116 GTTTCATATATTAAATGATATTGTTCAAAATGGAAGGCGCAAAAGCGAGTTCAGCAATTT 57
   |||||||
QY 1653 tcaaacgtatgatacttaataaataactataaataaataatataatttc 1705
   |||||||
Db 56 TCAAACTGATGATCAATTAATAAATACTATTAATTAATAATATATATTTTC 4
   |||||||

```

RESULT 8

AAFS5043
 ID AAF55043 standard; DNA; 1316 BP.

XX AAF55043;
 AC
 XX

15-MAY-2001 (first entry)

XX Nucleotide sequence of a calcium channel transport polypeptide.
 DE
 XX
 XX Calcium channel transport polypeptide; calcium trafficking;
 KW neural disorder; HIV-induced dementia; immune system disorder;
 KW rheumatoid arthritis; muscular disorder; muscle contractile dysfunction;
 KW reproductive disorder; gastrointestinal disorder; pulmonary disorder;
 KW cardiovascular disorder; arrhythmia; renal disorder;
 KW proliferative disorder; cancer; lung carcinoma; breast cancer; ss.

OS Homo sapiens.

EH Key Location/Qualifiers

FT CDS 36..341

FT /tag= a

XX /product= "calcium channel transport polypeptide"

PF 27-JUL-2000; 2000MO-US20392.

PR 28-JUL-1999; 99US-0145958.

PR 18-AUG-1999; 99US-0149446.

PR 14-MAR-2000; 2000US-0189064.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, NI J, Shi Y;

DR WPI; 2001-138604/14.

XX P-PSDB; AAB67466.

PT New isolated nucleic acid useful for diagnosing, detecting, or treating

PT or preventing diseases associated with anomalies in calcium trafficking

PT across the plasma membrane -

XX Claim 1; Page 254; 259pp; English.

PS The present sequence encodes a calcium channel transport polypeptide.

XX The polynucleotides, polypeptides, and antibodies are useful for

CC preventing, treating, or ameliorating diseases associated with anomalies

CC in calcium trafficking across the plasma membrane. They are used to

CC diagnose, detect and treat or prevent diseases or conditions such as

CC neural disorders (e.g. HIV-induced dementia), immune system disorders

CC (e.g. Rheumatoid arthritis), muscular disorders (e.g. muscle contractile

CC dysfunction), reproductive disorders, gastrointestinal disorders, renal

CC pulmonary disorders, cardiovascular disorders (e.g. arrhythmias), renal

CC disorders, proliferative disorders, and/or cancerous diseases and

CC conditions (e.g. lung carcinoma or breast cancer).

XX Sequence 1316 BP; 458 A; 227 C; 237 G; 394 T; 0 other;

Query Match 43.3%; Score 738.6; DB 22; Length 1316;
 Best Local Similarity 98.7%; Pred. No. 7.5e-159;
 Matches 755; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

```

QY 933 aaaaaaacctcaagaaagcacaataatgctacacaaagtgcacaatttta 992
   |||||||
Db 526 aaaaaaacctcaagaaagcacaataatgctacacaaagtgcacaatttta 585
   |||||||
QY 993 agaagcctcatgagctgtagtggggaattcagaaaaaaatacaggaagaaca 1052
   |||||||
Db 586 agaagcctcatgagctgtagtggggaattcagaaaaaaatacaggaagaaca 645
   |||||||
QY 1053 caccagaagggttttttccctacaccagcaagaacatatatagatataatctca 1112
   |||||||
Db 646 caccagaagggttttttccctacaccagcaagaacatatatagatataatctca 705
   |||||||
QY 1113 attataattatgcatatttgcattttatctcaaaataacttgggggaacatgta 1172
   |||||||

```

```

Dh 706 attataattatgcatttaatttcattttatttcaaaatcatttctgttgggacatgtaac 765
Qy 1173 cctctgagcaatcgaatatttttgggaatcctcttaaaagtacaattatcaatlaaa 1232
Dh 766 cctctgagcaatcgaatatttttgggaatcctcttaaaagtacaattatcaatlaaa 825
Qy 1233 ttaactagtagaataatgtaattcagaacaaagaataatcagaatattgagttgcttg 1292
Dh 826 ttaactagtagaataatgtaattcagaacaaagaataatcagaatattgagttgcttg 885
Qy 1293 ctggtgttgaagcaccatgtagatcattcacaagtctgcaaaagtcaaaacatactgt 1352
Dh 886 ctggtgttgaagcaccatgtagatcattcacaagtctgcaaaagtcaaaacatactgt 945
Qy 1353 acatcacaacccaagaatacacaataatccagaatagagcccatataatgatttaata 1412
Dh 946 acatcacaacccaagaatacacaataatccagaatagagcccatataatgatttaata 1005
Qy 1413 catgtagcttttgcacataataaagccattggaacacggaagaattagataactaataacat 1472
Dh 1006 catgtagcttttgcacataataaagccattggaacacggaagaattagataactaataacat 1065
Qy 1473 tgaactatcttttgaataacagcactcaaatgtagttagcttttccatggtggaat 1532
Dh 1066 tgaactatcttttgaataacagcactcaaatgtagttagcttttccatggtggaat 1125
Qy 1533 ttttaattacttttcccttctgttaatttctctctgtatatttaacaataagctgtgata 1592
Dh 1126 ttttaattacttttcccttctgttaatttctctctgtatatttaacaataagctgtgata 1184
Qy 1593 gtttacaatatataaagaatttgcataatttgaaggcgaagccaggtccgaacttt 1652
Dh 1185 gtttacaatatataaagaatttgcataatttgaaggcgaagccaggtccgaacttt 1244
Qy 1653 tcaaatgtatgtacatttaataaataactataaataaataa 1697
Dh 1245 tcaaatgtatgtacatttaataaataactataaataaataa 1289

```

RESULT 9

```

ABA08976
ID ABA08976 standard; cDNA; 1353 BP.
XX
AC ABA08976;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human secreted protein homologue-encoding cDNA, SEQ ID NO:752.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolytic; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritis; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antitumor; ss.
XX
OS Homo sapiens.
XX
XX
XX WO200157188-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US03800.
XX
XX 03-FEB-2000; 2000US-0496914.
XX
XX 27-APR-2000; 2000US-0560875.
XX

```

```

XX
PA (HYSE-) HYSEQ INC.
PI Tang YL, Liu C, Drmanac RT;
XX
XX WPI: 2001-457740/49.
XX P-PSDB: ABB11732.
XX
PS Human proteins and DNA encoding sequences useful for preventing,
PS treating or ameliorating a medical condition in a mammalian subject
PS e.g. arthritis and cancer -
PS
PS Claim 1: Page 677; 1963pp; English.

```

Sequences ABB10961-ABB12330 represent 1350 novel human polypeptides, and sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haematopoiesis regulatory activity; tissue growth activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to the use in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a cDNA encoding a novel human polypeptide of the invention.

Sequence 1353 BP; 349 A; 298 C; 315 G; 391 T; 0 other;

Query Match 40.3%; Score 687; DB 22; Length 1353;
Best Local Similarity 100.0%; Pred. No. 4.5e-147;
Matches 687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 8 gaagaagctttccacacagcgtgtatgcacgtgtgagtcggcgccatgtagaagaa 67
Dh 7 gaggaagctttccacacagcgtgtatgcacgtgtgagtcggcgccatgtagaagaa 66
Qy 68 atgatcgaggaatacacaagcctaatttgtaataataaagaagaataatttccct 127
Dh 67 atgatcgaggaatacacaagcctaatttgtaataataaagaagaataatttccct 126
Qy 128 tgaactatctccgtaaaagccattgtgctcctctcttgagggtgacaggtgttacaatcaa 187
Dh 127 tgaactatctccgtaaaagccattgtgctcctctcttgagggtgacaggtgttacaatcaa 186
Qy 188 tggccttgaaggagccttgagtcgaagttggagagccttggagagattcagatgctcattc 247

```

Db	187	tgccctttagagagcgtcttgatgcaaglttgaggcccttgggagatttcagatgtgtcaac	246
QY	248	tggttttatctctccctctcgaigttaataccctcatactctgtagagaacttgg	307
Db	247	tggttttatctctccctctcgaigttaataccctcatactctgtagagaacttgg	306
QY	308	ctgcagccatctcctgtgctcatcgtttgttggtccacatgtctgacacataatactgtagctg	367
Db	307	ctgcagccatctcctgtgctcatcgtttgttggtccacatgtctgacacataatactgtagctg	366
QY	368	gttaatgaaacttgaaatccctcagtgaaagatgcctcttgtagaattcctatacccaatagct	427
Db	367	gttaatgaaacttgaaatccctcagtgaaagatgcctcttgtagaattcctatacccaatagct	426
QY	428	caaatctggaagccgaagaagtgtcgtgtgcttttccatccccagttggcagctctctacc	487
Db	427	caaatctggaagccgaagaagtgtcgtgtgcttttccatccccagttggcagctctctacc	486
QY	488	tgaatggagactatccacagacagaatgtaggcagacacagaacctgtgtgatatggcttgg	547
Db	487	tgaatggagactatccacagacagaatgtaggcagacacagaacctgtgtgatatggcttgg	546
QY	548	tatatgacacaaagctactctccctcgcagacattgtagactaagtggagaccgtgtatgtgatt	607
Db	547	tatatgacacaaagctactctccctcgcagacattgtagactaagtggagaccgtgtatgtgatt	606
QY	608	atcacgtacatgaaatcacgtgtgttcaattcctactctgtagctggaatcctgttggagagca	667
Db	607	atcacgtacatgaaatcacgtgtgttcaattcctactctgtagctggaatcctgttggagagca	666
QY	668	tcataggtggccatgtctccagacaggt	694
Db	667	tcataggtggccatgtctccagacaggt	693

RESULT	10						
ID	AAI59215						
ID	AAI59215 standard; cDNA; 1375 BP.						
XX							
AC	AAI59215;						
XX							
DT	22-OCT-2001 (first entry)						
XX							
DE	Human polynucleotide SEQ ID NO 1418.						
XX							
KW	Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;						
KW	peripheral nervous system; neuropathy; central nervous system; CNS;						
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;						
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;						
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;						
KX	leukaemia; ss.						
XX							
OS	Homo sapiens.						
XX							
PN	WO200153312-A1.						
XX							
PD	26-JUL-2001.						
XX							
PF	26-DEC-2000; 2000WO-US34263.						
XX							
PR	21-JAN-2000; 2000US-0488725.						
PR	25-APR-2000; 2000US-0552317.						
PR	09-JUL-2000; 2000US-0598042.						
PR	19-JUL-2000; 2000US-0620312.						
PR	03-AUG-2000; 2000US-0653450.						
PR	14-SEP-2000; 2000US-0662191.						
PR	19-OCT-2000; 2000US-0693036.						
PR	29-NOV-2000; 2000US-072344.						
XX							
PA	(HYSE)- HYSOQ INC.						
XX							
Tang YT,	Liu C,	Asundi V,	Chen R,	Ma Y,	Qian XB,	Ren F,	Wang D
PI	Wang J,	Wang Z,	Wehrman T,	Xu C,	Xue AJ,	Yang Y,	Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR P-PSDB: AAM40059.
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
PS Claim 1: SEQ ID NO 1418; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AAM42213) with neurotrophic,
XX immunosuppressant and cyostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localized neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX Sequence 1375 BP; 354 A; 329 C; 324 G; 368 T; 0 other;
SQ

	Query Match	Similarity	23.4%	Score 399	DB Z2	Length 157	
	best Local	Similarity	77.0%	Pred. No. 2.e-81			
	Matches	486	Conservative	0	Mismatches	145	Indels
							Gaps
							0
QY	64	agaatattgatgagagatcaataaagcttaatttaatttaataaagaagaattttct	123				
DB	165	agagatgctgggtcagagatcaaaaacaatttagtgtaacttaaggaaagaacaatttcc	224				
QY	124	cctttgaactatctctcgttaaagccattgctccctcttgggggtcagaattttaa	183				
DB	225	ctctttgacctctctgataacagtcatattgtgctctcatttggagatcaactgtcaac	284				
QY	184	tcaatgaccttggggagctcttgaatcaagtgtgaagaccttggagattcagaatgct	243				
DB	285	tcaatgaccttccagagacctcttggtaacgtgtgtaacctgtgagatccagatcctt	344				
QY	244	catctggttttatctcttcctctcctcaatgttatataacctcaatactcctagaac	303				
DB	345	cagactggttttctctcaatctcttgcgtgtgctacataccttcatattgtctggagac	404				
QY	304	tttgtgtgacgacatctcgtgcatcgttgcctggtgcacatgcttggacaataactgga	363				
DB	405	ttaacagcttcaatacctgtgcacatcgtcgtgcgtgcacatccttgaacaatgaactgc	464				
QY	364	tctgtatgaatgaactggaatcctcctcagtgaaagtgtgcctcttgaagatctcatccacta	423				
DB	465	tctgaacaatgaactatgggtgccttcagccaagatgacatcttgaagatctcatccactg	524				
QY	424	gactcaaatctgggccaagagaagtgcgtgcgttgttccatccccaagtggagctttt	483				
DB	525	gactcaaaactgagcgccagagaaggtgcgtgcgttgttccatcccttaagtgagacctt	584				
QY	484	caactgaatggagactatccacagcaaaaagtgtagagcagacaacgaacctgttggatgagc	543				
DB	585	caactgaatggagaccttccccaacaagaatgtagcagacaatggagacctgtgtgatatgc	644				
QY	544	tgggtatgatcaataagcttactcctctggacacattgtgactaagtgggaacctggtatg	603				
DB	645	tgggtatgatcaagaatctctctctcaacacacatcgtgactgaagtgggaacctgtgatg	704				
QY	604	gataatacgtacatgaataatcagtggttcaattctactacttgaacttggaaatgctgttggga	663				
DB	705	gactcctaatacagacttcaagtggtcctaattgttatctatgctgagaaatgagtgtggga	764				

QY 664 ggcacataggtgacatgctcagacagct 694
 DB 765 ggcacataggtgacatgctcagacagct 795

RESULT 11

ID AAD09561 standard; cDNA; 1986 BP.

AC AAD09561;

DT 10-SEP-2001 (first entry)

DE Human transporter and ion channel-10 (TRICH-10) cDNA.

Human; transporter and ion channel-10; TRICH-10; cystic fibrosis; mood; gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy; hypertension; angina; neurological disorder; asthma; bipolar disorder; dementia; depression; Alzheimer's disease; epilepsy; vaccine; arrhythmia; Pick's disease; Ischaemic cerebrovascular disease; AIDS; anxiety; stroke; Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy; demyelinating disease; mental disorder; Schizophrenia; polyomyelitis; muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease; dermatomyositis; diabetes mellitus; immunological disorder; psoriasis; rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus; sickle cell anaemia; Wilson's disease; infertility; Cushing's disease; scleroderma; pulmonary artery stenosis; neutropenic; Addison's disease; malabsorption syndrome; hypercholesterolaemia; cancer; ss.

OS Homo sapiens.

Key Location/Qualifiers
 CDS 248..1909

FT /tag=a
 FT /product="Human TRICH-10 protein"

PN MO200146258-A2.

PD 28-JUN-2001.

PF 22-DEC-2000; 2000MO-US35095.

PR 23-DEC-1999; 98US-0172000.

PR 14-JAN-2000; 2000US-0176083.

PR 21-JAN-2000; 2000US-0177332.

PR 28-JAN-2000; 2000US-0178572.

PR 02-FEB-2000; 2000US-0179758.

PR 10-FEB-2000; 2000US-0181625.

PA (INCY-) INCYTE GENOMICS INC.

PI Baughn MR, Burford N, Au-Young J, Lu DM, Yang J, Reddy R, Lal P;

PI Hillman JL, Azimzai Y, Yue H, Nguyen DB, Yao MG, Gandhi AR;

PI Tang YT, Khan FA;

XX WPI; 2001-418042/44.

XX P-PSDB; AAE04897.

PT Novel human transporter and ion channel proteins useful for treating

PT and preventing transport, neurological, muscle and immunological

XX disorders

XX Claim 5; Page 147; 160pp; English.

XX The present sequence is transporter and ion channel-10 (TRICH-10) cDNA.

XX condition associated with decreased expression of functional TRICH,

XX such as transport disorder including amyotrophic lateral sclerosis,

XX cystic fibrosis, Becker's muscular dystrophy, Charcot-Marie Tooth

XX disease, Duchenne muscular dystrophy, angina and hypertension,

XX neurological disorders including Alzheimer's disease, amnesia, bipolar

XX disorder, dementia, depression, epilepsy, ischaemic cerebrovascular

CC disease, stroke, cerebral neoplasms, Pick's disease, Huntington's
 CC disease and Parkinson's disease, demyelinating diseases, mental disorders
 CC including mood, anxiety, schizophrenia and seasonal affective disorder,
 CC muscle disorder including cardiomyopathy, myocarditis, polyomyelitis,
 CC dermatomyositis, arrhythmias and asthma and immunological disorders
 CC including AIDS, adult respiratory distress syndrome (ARDS), allergies,
 CC anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's
 CC syndrome, systemic lupus erythematosus and other diseases including
 CC sickle cell anaemia, Wilson's disease, cataracts, infertility, pulmonary
 CC artery stenosis, Grave's disease, Cushing's disease, Addison's disease,
 CC glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers
 CC psoriasis and viral, bacterial, fungal, helminthic and protozoal
 CC infections. TRICH DNA is useful in gene therapy and in diagnostic
 CC purposes.
 XX

SO Sequence 1986 BP; 505 A; 496 C; 457 G; 528 T; 0 other;

Query Match 23.4%; Score 399; DB 22; Length 1986;
 Best Local Similarity 77.0%; Pred. No. 2,5e-81;

Matches 486; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 64 agaatggtgagagatacaacagcttaattgaattaaagaatatttct 123
 DB 125 agagagtggtgagagatacaacagcttaattgaattaaagaatatttct 184
 QY 124 ccttgaactatctcgttaagcatttgccctcttggtgggtacgtgttcaaa 183
 DB 185 ccttgaactatctcgttaagcatttgccctcttggtgggtacgtgttcaaa 244
 QY 184 tcaatgaccttgagagaccttgagaaagtgtgagagaccttgagaaagtgtgag 243
 DB 245 tcaatgaccttgagagaccttgagaaagtgtgagagaccttgagaaagtgtgag 304
 QY 244 catcgttttattctcctcctcctcctcctcctcctcctcctcctcctcctcct 303
 DB 305 cagactgttttattctcctcctcctcctcctcctcctcctcctcctcctcctcct 364
 QY 304 ttctgtcagcattcctcctcctcctcctcctcctcctcctcctcctcctcctcct 363
 DB 365 ttctgtcagcattcctcctcctcctcctcctcctcctcctcctcctcctcctcct 424
 QY 364 tctgttaagaaacttgagaaaccttgagaaaccttgagaaaccttgagaaaccttgag 423
 DB 425 tctgttaagaaacttgagaaaccttgagaaaccttgagaaaccttgagaaaccttgag 484
 QY 424 gactcaaatctgag 483
 DB 485 gactcaaatctgag 544
 QY 484 cactcaaatctgag 543
 DB 545 cactcaaatctgag 604
 QY 544 tgggtatgatcaaaagactcctcctcctcctcctcctcctcctcctcctcctcct 603
 DB 605 tgggtatgatcaaaagactcctcctcctcctcctcctcctcctcctcctcctcct 664
 QY 604 gattatcagctcaaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat 663
 DB 665 gattatcagctcaaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat 724
 QY 664 ggcacataggtgacatgctcagacagct 694
 DB 725 ggcacataggtgacatgctcagacagct 795

RESULT 12

ID AAC61892 standard; cDNA; 2027 BP.

AC AAC61892;

XX

Tue, Jul 2 10:13:50 2002

us-09-674-235-19.rng

Page 15

06-MAR-2001 (first entry)
cDNA encoding a human secreted protein.
Secreted protein; platelet disorder; stem cell disorder; osteoporosis;
secreted protein; burn; incision; ulcer; periodontal disease; neuropathy;
neovascular disease; bone growth; cosmetic plastic surgery;
neovascular disease; bone growth; cosmetic plastic surgery;
gut protection; gut regeneration; fibrosis; cancer;
bone marrow transplantation; ss.
Homo sapiens.
Location/Qualifiers
Key 270..1289
CDS /tag="a
/product="secreted protein"
MO200061755-A2.
19-OCT-2000.
10-APR-2000; 2000WO-US09555.
09-APR-1999; 99US-0128574.
20-AUG-1999; 99US-0150054.
(CHIR) CHIRON CORP.
Garcia PD;
WPI: 2000-665133/64.
P-SDB: AAB19406.
Novel secreted human proteins useful for stimulating blood cell
generation in patients receiving cancer chemotherapy, treating bone
marrow transplantation patients and for healing fractured bones
Claim 2: Page 71-72; 74pp; English.
AC61879-93 encode secreted human proteins. The secreted proteins are
useful in assays to determine their biological activities. The proteins
can also be used as biomarkers to identify tissues or cell types which
express the proteins. The polynucleotide molecules can be used as
biomarkers for tissues or chromosomes and to elicit immune responses.
The proteins and antibodies are useful in diagnosis and treatment of
diseases associated with altered expression of these proteins. The
proteins are also useful for prevention or treatment of platelet
disorders, stem cell disorders, osteoporosis or osteoarthritis, burns,
incisions, ulcers, periodontal diseases, central and peripheral nervous
system diseases and/or bone growth in cosmetic plastic surgery. The
proteins are also useful for gut protection or regeneration, for the
treatment of lung or liver fibrosis, for stimulating blood cell
generation in patients receiving cancer chemotherapy and for treatment
of bone marrow transplantation patients.
Sequence 2027 BP; 530 A; 501 C; 459 G; 537 T; 0 other;
Query Match 23.28; Score 395.8; DB 21; Length 2027;
Best Local Similarity 76.78; Pred. No. 1.3e-80;
Matches 484; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
QY 64 agaatgagcagagatcaatcaagaagcattgaatgaataaagaagaattttct 123
DB 147 agagaggggggtaagatacaaacacattgagcagcagcagcagcagcagc 206
QY 124 cccttgactatctccgtaagacatgctcctcctgggggagcagcagcagcagc 266
DB 207 cttctgacacctcctgagacagcagcagcagcagcagcagcagcagcagc 243
QY 184 tcaatgagccttgagagcagcagcagcagcagcagcagcagcagcagcagc 243

DB 267 tcaatgagccttgagagcagcagcagcagcagcagcagcagcagcagcagc 326
QY 244 catctggttttattctccctcctcctcctcctcctcctcctcctcctcctcct 366
DB 327 cagactggtttctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 363
QY 304 ttgtctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 446
DB 387 ttacgacatgacatgacatgacatgacatgacatgacatgacatgacatgacat 423
QY 364 tctgagacatgacatgacatgacatgacatgacatgacatgacatgacatgacat 506
DB 447 tctgagacatgacatgacatgacatgacatgacatgacatgacatgacatgacat 483
QY 424 gactcaaatctgagagcagagagagagagagagagagagagagagagagagagag 566
DB 507 gactcaaatctgagagcagagagagagagagagagagagagagagagagagagag 543
QY 484 cactgagatgagagcagagagagagagagagagagagagagagagagagagag 626
DB 567 cactgagatgagagcagagagagagagagagagagagagagagagagagagag 603
QY 544 tggatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 686
DB 627 tggatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 663
QY 604 gattatcagatcagatcagatcagatcagatcagatcagatcagatcagatcagat 746
DB 687 gactcaaatctgagagcagagagagagagagagagagagagagagagagagagag 694
QY 664 gactcaaatctgagagcagagagagagagagagagagagagagagagagagagag 777
DB 747 gactcaaatctgagagcagagagagagagagagagagagagagagagagagagag
RESULT 13
AAC85823 standard; DNA; 1977 BP.
ID AAC85823;
AC AAC85823;
XX 06-AUG-2001 (first entry)
DE hOAT4 DNA.
XX Human: organic anion transporter; hOAT; liver; kidney;
KW membrane protein; transport; organic anion; splice variant; ds.
XX Homo sapiens.
OS Homo sapiens.
FH Key 236..1900
FT CDS /tag="a
FT /product="hOAT4"
PN WO200104283-A2.
XX 18-JAN-2001.
PD 12-JUL-2000; 2000WO-US18980.
PE 12-JUL-1999; 99US-0143771.
PR (META-) METABASIS THERAPEUTICS INC.
PA Sun W;
PI Sun W;
PI WPI: 2001-367057/38.
DR P-SDB: AAB47275.
XX Nucleic acids encoding human organic anion transporter polypeptides,
XX useful in gene therapy procedures -

PS Claim 1; Fig 5; 95pp; English.

CC The sequences given in AAC85819-24 encode human organic anion
CC transporter (HOAT) polypeptides. HOAT polypeptides are preferentially
CC expressed in the liver and kidneys of humans. OAT's are membrane
CC proteins that facilitate the transport of organic anions across the
CC cell membrane. The mechanism of transport of organic anions across the
CC HOAT2A and HOAT2B are thought to be splice variants of a secondary
CC identical except at the C-terminal end. HOAT proteins as they are
CC of diseases associated with inappropriate HOAT expression.

Sequence 1977 BP; 502 A; 491 G; 450 G; 534 T; 0 other;

Query Match

Best Local Similarity 22.2%; Score 378; DB 22; Length 1977;
Matches 481; Conservative 0; Mismatches 150; Indels 3; Gaps 1;

```

OY 64 agaaatgatgcgagatcaataaagcttaattgaataaagaatatttct 123
DB 113 agagagtggttcagatcaataaacaattagtgacttagggaagaacatttcc 172
OY 124 ccccttaactatctcgtgaagccatgtgctcctcttgagggtcagtggtcaca 183
DB 173 ctcttgaaacctcctggaatacagctatttgcctcactgaagatcaactgttcaac 232
OY 184 tcaatgaccttgagagatccttgagatcagtgagagcttgaggagattcagatgct 243
DB 233 tcaatgaccttgagagatccttgagatcagtgagagcttgaggagattcagatgct 292
OY 244 catctgttttattctcctcctcctcctcctcctcctcctcctcctcctcctcct 303
DB 293 cagactgttttctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 352
OY 304 ttgtgtagcagcattcctcctcctcctcctcctcctcctcctcctcctcctcct 363
DB 353 ttcactgattcattacactcctcctcctcctcctcctcctcctcctcctcctcct 412
OY 364 tctgtaataaactggaatcctcctcctcctcctcctcctcctcctcctcctcct 423
DB 413 tctgtaataaactggaatcctcctcctcctcctcctcctcctcctcctcctcct 472
OY 424 gactcaatctgagagcagaagatgctcctcctcctcctcctcctcctcctcctcct 483
DB 473 gactcaatctgagagcagaagatgctcctcctcctcctcctcctcctcctcctcct 532
OY 484 cactgaatctgagagcagaagatgctcctcctcctcctcctcctcctcctcctcct 543
DB 533 cactgaatctgagagcagaagatgctcctcctcctcctcctcctcctcctcctcct 592
OY 544 tgggataatgtaaaactcctcctcctcctcctcctcctcctcctcctcctcctcct 600
DB 593 tgggataatgtaaaactcctcctcctcctcctcctcctcctcctcctcctcctcct 652
OY 601 tggatatacagtcagtaaatcagtggttcaatcctcctcctcctcctcctcctcct 660
DB 653 tggatatacagtcagtaaatcagtggttcaatcctcctcctcctcctcctcctcct 712
OY 661 gtaggacatcatagtgagcagtgctcagacaggt 694
DB 713 gtaggacatcatagtgagcagtgctcagacaggt 746

```

RESULT 14
AAI61001
ID AAI61001 standard; cDNA; 1132 BP.
AC AAI61001;
XX
XX
DT 22-OCT-2001 (first entry)

DE Human polynucleotide SEQ ID NO 4990.

XX Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukemia; ss.

OS Homo sapiens.

PN WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000MO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX P-PSDB; AAM41845.

XX Novel nucleic acids and polypeptides, useful for treating disorders

XX such as central nervous system injuries.

XX Claim 1; SEQ ID NO 4990; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and

XX the encoded polypeptides (AAM38642-AAM42213) with nocotropic,

XX immunosuppressant and cytostatic activity. The polynucleotides are useful

XX in gene therapy. A composition containing a polypeptide or polynucleotide

XX of the invention may be used to treat diseases of the peripheral nervous

XX system, such as peripheral nervous injuries, peripheral neuropathy and

XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

XX utilisation of the activities such as: immune system suppression,

XX activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic

XX assays for receptor activity, cancer diagnosis and therapy, drug screening,

XX C.N.S. disorders.

XX Note: The sequence data for this patent did not form part of the printed

XX specification.

SO Sequence 1132 BP; 274 A; 265 C; 285 G; 305 T; 3 other;

Query Match

Best Local Similarity 13.7%; Score 232.8; DB 22; Length 1132;
Matches 264; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

OY 379 ggaatcctcagtgagagcctcttgagaatcctacccacagacacaaatctgagg 438

DB 27 ggggacctcagcaagaagacacctcttgagaatcctacccacagacacaaatctgagg 86

OY 439 ccagaagatgctgctcttgctcaccacccagtgagcactcctcaccatggaatggagct 498

DB 87 ccagaagatgctgctcttgctcaccacccagtgagcactcctcaccatggaatggagct 146

OY 499 atccacagcacaagtggagcagacacagaccctctgtgagtgctggttatatgacaa 558

...

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
July 1, 2002, 15:41:50 ; Search time 3205.87 Seconds
(without alignments)
1129.505 Million cell updates/sec

Run on: July 1, 2002, 15:41:50 ; Search time 3205.87 Seconds
(without alignments)
1129.505 Million cell updates/sec

Title: US-09-674-235-19
1705
1 aagactgaggaagctctt.....aaattaaaaattatttc 1705

Sequence: 1705
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Scoring table: 1797656 seqs, 10463268293 residues
3595312

Searched: 1797656 seqs, 10463268293 residues
3595312

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hgt_hum:*
31: em_hgt_in:*
32: em_hgt_other:*
33: em_hgtgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	801.8	47.0	2684	6	AX074150
2	774.8	45.4	164809	2	AC090778
3	774.8	45.4	169776	2	AP001880
4	762.8	44.7	166848	2	AP002367
5	761.2	44.6	170682	2	AC073256
6	759.6	44.6	116273	2	AP000596
7	588	34.5	160945	2	AP003420
8	588	34.5	166848	2	AP002367
9	586.4	34.4	116273	2	AP000596
10	400	23.5	1566	6	BC022379
11	399	23.4	1986	6	AX179749
12	380.6	22.3	1846	6	AB062418
13	378	22.2	1977	6	AX074149
14	340.8	20.0	160945	2	AP003420
15	340.8	20.0	164809	2	AC090778
16	340.8	20.0	169776	2	AP001880
17	324.6	19.0	185404	2	AC025409
18	320.8	18.8	179934	2	AP000810
19	320.8	18.8	203598	2	AC090250
20	320.8	18.8	210723	2	AP001858
21	301.4	17.7	2016	10	BC014805
22	301.4	17.7	2878	10	MMU132857
23	300.4	17.6	1911	10	RNO132859
24	300	17.6	2597	10	RNUT1R
25	298.2	17.5	2253	9	AK057654
26	293.8	17.2	1995	10	AB051836
27	290.6	17.0	1964	10	BC021384
28	281.2	16.5	2079	10	RNO132858
29	273.2	16.0	170682	2	AC073256
30	250.4	14.7	207064	2	AC106680
31	241.6	14.2	181234	6	AX103148
32	218.4	12.8	2865	6	AX191489
33	208.8	12.2	1737	6	AX191489
34	208.8	12.2	2455	6	AX230568
35	193	11.3	1779	10	AB005451
36	189.2	11.1	3258	9	AB050269
37	188	11.0	79141	2	AC091122
38	180.8	10.6	2516	6	AK05737
39	180.8	10.6	1653	6	AX191610
40	180.8	10.6	2047	6	AX230572
41	180.8	10.6	2194	6	AX191620
42	180.8	10.6	2210	6	E54995
43	180.8	10.6	2210	6	AB026116
44	170	10.0	119151	2	AP000655
45	170	10.0	120628	9	AC012153

ALIGNMENTS

RESULT	1	DNA	linear	PAT 06-FEB-2001
AX074150	AX074150	2684 bp		
LOCUS	Sequence 6 from Patent WO0104283.			
DEFINITION	AX074150			
ACCESSION	AX074150.1	GI:12710362		
VERSION				
KEYWORDS				
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	1 (Dases 1 to 2684)			
AUTHORS	Sun, W.			
TITLE	Organic anion transporter genes and proteins			
JOURNAL	Patent: WO 0104283-A 6 18-JAN-2001.			
FEATURES	Metabasis Therapeutics Inc. (US)			
source	1. 2684			
BASE COUNT	800 a 533 c 541 g 810 t			
ORIGIN	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			

Query Match 47.0% Score 801.8; DB 6; Length 2684;
 Best Local Similarity 66.2% Pred. No. 1.3e-146;
 Matches 1693; Conservative 0; Mismatches 12; Indels 854; Gaps 3;

QY 1 aagaactgaggaagctcttcacacagcgtatattgcaactgtygaagtcgggagccatg 60
 Db 14 AAGAAGCTAGAGAAAGCTTTCCACACGCGTGTATGCTACTGTGGAGTCCGGCCCATG 73
 QY 61 atgagaatgatgagagatcaatacaagcttaattgataataaaggaaatatt 120
 Db 74 ATGAGAAATGATGTCGAGAGATACATACACCTTAATTGATTAATAAAGAAATATT 133
 QY 121 tctcccttgaactatctccgtaagccatgtgctctcttgggggctcagytca 180
 Db 134 TCTCCCTTGAACCTTATCTCCGTAAGCCATGTGTGCTCTCTTGGGGGTACGCTTCA 193
 QY 181 caatcaatggccttgaagagccttgagcaagttggaagccttggagattcagatg 240
 Db 194 CAATCAATGGCCTTGAGAGCTTGAAGTCAAGTTGGAGGCTTGGAGATTTCAGATG 253
 QY 241 ctcaactggttttattctctccctctctcctcattgtatataatccatatactag 300
 Db 254 CTTCATGCTGTTTATTCTCTCCCTCTCTCATGTTAATCCCTCATATACGCTAGAG 313
 QY 301 aacttctgcaagccatctcctgctcactgctggtggtccacatgctgagcaataact 360
 Db 314 AACTTCTGTCAGCCATCTCGTCATCGTTCGCTGCTGCTCAATGCTGACATATACT 373
 QY 361 ggaatcgtgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 420
 Db 374 GGATCTGTGATGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACT 433
 QY 421 ctgaactcaatctgagcagcagagagtgctgctgctgctgctgctgctgctgctgct 480
 Db 434 CTGAGCTCAATCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 493
 QY 481 cttaactgaatgagactatccacagcaagtgagagagagagagagagagagagag 540
 Db 494 CTTCACCTGAATGAGACTATCCACAGCAAGTGGAGGAGCAACAGAACCTGTGTGAT 553
 QY 541 ggcctggatataatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 600
 Db 554 GCGCTGGATATGATCAAGCTACTCTCTCTGACACTGTGACTAAGTGGAGCTGTATG 613
 QY 601 tctgattatcagctgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 660
 Db 614 TGTGATTATCAGTCACTGAATCACTGCTTCAATCTCTTCTGACTGGAATGCTGCTG 673
 QY 661 ggaagcactaagtgagcactgctcagac----- 690
 Db 674 GGAAGCACTAAGTGGCCATGCTCAGACAGTTTGGGCAAGATTTATTCTCAGATGG 733
 QY 691 ----- 690
 Db 734 TGTGCTCCAGCTTGCCATTACTACACTGCGCTGCCCTTCCGCTCCACCTTCCCTGTT 793
 QY 691 ----- 690
 Db 794 TACTGTACTAAGCTTCTTGGCAGGTTTTCTTCATGATCATATATATCAATAATATTCT 853
 QY 691 ----- 690
 Db 854 TTGCCCCATTACTAGATGATAGGCCCACTTAAGCCCTGTAGTAATATTGTCATCT 913
 QY 691 ----- 690
 Db 914 GGTGCCCTTAGATTGGACAGATATCTGGAGGCTTGGCTTATGCTTCCGAGACTGG 973
 QY 691 ----- 973
 Db 974 CAACCTGACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1033

QY 697 ctgtggaatctgctcgttgggtgtaatacaaccaataaactagatgagggcttaagga 756
 Db 1034 CTGTGTGAATCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 756
 QY 757 cttaaaaaagttgcaagcaacaatggaataaagaatgctgaagaaacccctgaacataag 816
 Db 1094 CTTAGAAAAGTTGCAAGCAGCAATGGAATTAAGAAATCTGMAAACCCTGAACATAGAG 1153
 QY 817 gttgtaagatccaccatgcaagagagagagctggaatgcaagcaagcaagcaagcaag 876
 Db 1154 GTTGAAGATCCACATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1213
 QY 877 gactgttccgcaacccagtaagtgtaagagatctgatacttgatatttga----- 932
 Db 1214 GACTTGTCCGCAACCCCGATATGCTTAAGAGATCTGATCTGCTGCTGCTGCTGCTGCT 932
 QY 933 ----- 932
 Db 1274 GCAAAACAAATACCTTTTATGTAACAATGATCAATCTTACATGAGGAGCAACATT 1333
 QY 933 ----- 932
 Db 1334 TTCTGTTGCAAGTACTTATGAGAGCTGCTGCTCATAGTTGATGATGCTTGTCTTTG 1393
 QY 933 ----- 932
 Db 1394 ACATTAATCATATGAGGCGCTGCAATTAAGCAATATTTGATGCTTCTGCTGCTGCT 932
 QY 933 ----- 932
 Db 1454 TCCATTTTGGCAACAACGTTTGGCCCAAGAAATGACAGACCCTGCTGCTGCTGCT 932
 QY 933 ----- 932
 Db 1514 TGTCTGGAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 932
 QY 933 ----- 932
 Db 1574 CTGATCCCACTGTTCTGAGGCAAGAGCTTCAAGATATTAAGGCTAGTAGATT 1633
 QY 933 ----- 932
 Db 1634 GGAGGAGACTGCTGCCCTCTTGATGACCTTAACGATATTTTACACTTCCCATG 1693
 QY 933 ----- 932
 Db 1694 ATCATTTATGAACTTCCCATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 932
 QY 933 ----- 932
 Db 1754 AAGAACTGCTTGGCTGACACCATCAAGATGTGAAATCAAAAAAATCTCAAG 1813
 QY 948 gaaaagcaataaaatgattgctacaaaagtgcaaatcttaagaaacctcactagag 1007
 Db 1814 GAAAAGCATAAATGATTGCTACCAAAAAGTACCAAAATTTTAAACCTCTTCATGAG 1873
 QY 1008 ctgattgtggggaatctcagaaaaaaatacagggaaagaaagcaagcaagagggctt 1067
 Db 1874 CTGATTGCTGGGAAATTCAGAAAAAATAACAGGAAAAAGAACACCAAGAGGCTTT 1933
 QY 1068 ttccctacaacagcaagacatatatagatacagtaactcaatlaataatgagca 1127
 Db 1934 TTTCCCTACAAACAGCAAGAAATATATATGATACATGATCTCAATATATATTTGCA 1993
 QY 1128 ttaattgcaattatlaataaataactgtyggagacatgtaactcttgaagcaactg 1187
 Db 1994 TTAATTTGCAATTTATTAATAATTAATTTGAGGAGATGTAATCTTGAAGCAATCTG 2053
 QY 1188 atatttttggagagtccttaaaaaagttacaaattatcaataaactagtagaag 1247
 Db 2054 ATATTTTTGGAGTCTTTTAAAAAGTTTACAAATTTTATCAATTAATTAATTAATTAAT 2113
 QY 1248 atgattcagaacaaagaataacagaaatagatgtygctggc-tggtgatagaagc 1306

```

|||||
Db 2114 ATGATTCAGAAAAGAAATACAGAAATTAGATGCTGCTGCTGATGAGC 2173
|||
Qy 1307 accatgctgattcacaagaattgcaaaagtcacaaatactgctacacacagaa 1366
|||
Db 2174 ACCATGTGATGATTCATTAAGTTCAAAAGCAAAACATCTGTACATCCACGAA 2233
|||
Qy 1367 atcaataataacagagaatagacactataaattgcatthaatacgtactttga 1426
|||
Db 2234 ATCAAAATTAATCCAGAAATAGACACTATATAATGATTAATCATGATCTTTGA 2293
|||
Qy 1427 cataataagccattggaacggaagattagatactataaatacattgactctcttg 1486
|||
Db 2294 CATATTAAGCCATTGGAACGGAAGATGATGATCTTAATACATGCTATCTTTG 2353
|||
Qy 1487 taataacagcactaaatgctgattactcttccatgctggaatttaattactttt 1546
|||
Db 2354 TAAATACAGTCACATAATGATGATGATCTTTCATGCTGGAATTAATTAATCTTTT 2413
|||
Qy 1547 ctctgtaatttctctctgtatattttaacaataagctgctgatagttacaataat 1606
|||
Db 2414 CTCTGTAATTTCTCTCTGATATTTTAACAATAATGCTGATAGTTTACATATTAAT 2473
|||
Qy 1607 aaagatattgtaaatgaaaggcgaagccaggttcagcaatttcaacactgtatga 1666
|||
Db 2474 AAAGATATGTTCAATTAAGGCAAGCCAGCTTCACCAATTTTCAACGTATGTA 2533
|||
Qy 1667 cattataaataactataaataaataattatcttc 1705
|||
Db 2534 CATTAATTAATTAATTAATTAATTAATTAATTTTC 2572
|||

```

```

RESULT 2
AC090778 164809 bp DNA linear HTG 27-JAN-2002
LOCUS Homo sapiens chromosome 11 clone CTD-3110P2 map 11, WORKING DRAFT
DEFINITION
AC090778.2 GI:14336491
ACCESSION AC090778
VERSION AC090778.2
KEYWORDS HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 164809)
Bairren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Bairn, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A.,
Cammarata, J., Campolongo, A., Choquet, Y., Colangelo, M., Collins, S.,
Collins, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
Gardina, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagob, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Karatas, A., Lacroque, K., Lamares, R., Landers, T.,
Lenczky, J., Levine, R., Liu, G., Maclean, C., McDonald, P.,
Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
McPherson, R., Meldrum, J., Menus, L., Mhova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Piere, N., Pollara, V., Raymond, C., Retta, R.,
Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roselli, M.,
Roy, A., Santos, R., Schauer, S., Stange-Thomann, N., Stojanovic, N.,
Suzguez, C., Spencer, B., Stange-Thomann, N., Testa, S., Theodor, J.,
Strauss, N., Subramanian, A., Talamas, J., Testa, S., Theodor, J.,
Travers, M., Travers, N., Triggillo, J., Vassiliou, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zaitoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
TITLE
Direct Submission
JOURNAL
Submitted (10-MAR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```

COMMENT

```

On Jun 11, 2001 this sequence version replaced gi:13270647.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIDR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: 112735
Center clone name: 3110_P2
-----
Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 157859 bases at least Q40
Consensus quality: 161528 bases at least Q20
Consensus quality: 162934 bases at least Q20
Insert size: 172000; agarose-fp
Insert size: 163409; sum-of-ctdgs
Quality coverage: 5.0 in Q20 bases; agarose-fp
Quality coverage: 5.2 in Q20 bases; sum-of-ctdgs
-----
NOTE: This is a 'working draft' sequence. It currently
consists of 15 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1
1069: contig of 1069 bp in length
1070 1169: gap of 100 bp
1170 2379: contig of 1210 bp in length
2380 2479: gap of 100 bp
2480 3440: contig of 961 bp in length
3441 3540: gap of 100 bp
3541 4880: contig of 1340 bp in length
4881 4980: gap of 100 bp
4981 6768: contig of 1788 bp in length
6769 6868: gap of 100 bp
6869 8318: contig of 1450 bp in length
8319 8418: gap of 100 bp
8419 10046: contig of 1628 bp in length
10047 10146: gap of 100 bp
10147 12455: contig of 2309 bp in length
12456 12555: gap of 100 bp
12556 15164: contig of 2609 bp in length
15165 15264: gap of 100 bp
15265 16648: contig of 1384 bp in length
16649 16748: gap of 100 bp
16749 29429: contig of 12681 bp in length
29430 29529: gap of 100 bp
29530 41226: contig of 11697 bp in length
41227 41326: gap of 100 bp
41327 57370: contig of 16044 bp in length
57371 57470: gap of 100 bp
57471 102968: contig of 45498 bp in length
102969 103068: gap of 100 bp
103069 164809: contig of 61741 bp in length.
Location/Qualifiers
1..164809 Homo sapiens
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="CTD-3110P2"
/clone_11b="CTD Human BAC"
1..1069
/note="assembly_fragment"
1170..2379
/note="assembly_fragment"

```

```

misc_feature      2480..3440
                  /note="assembly-fragment"
misc_feature      3541..4880
                  /note="assembly-fragment"
misc_feature      4981..6768
                  /note="assembly-fragment"
misc_feature      6869..8318
                  /note="assembly-fragment"
misc_feature      8419..10046
                  /note="assembly-fragment"
misc_feature      10147..12455
                  /note="assembly-fragment"
misc_feature      12556..15164
                  /note="assembly-fragment"
misc_feature      15265..16648
                  /note="assembly-fragment"
misc_feature      16749..29429
                  /note="assembly-fragment"
misc_feature      29530..41226
                  /note="assembly-fragment"
misc_feature      41327..57370
                  /note="assembly-fragment"
misc_feature      57471..102968
                  /note="assembly-fragment"
misc_feature      103069..164809
                  /note="assembly-fragment"
BASE COUNT      50889 a 32076 c 31689 g 48754 t 1401 others
ORIGIN

```

```

Query Match      45.4%; Score 774.8; DB 2; Length 164809;
Best Local Similarity 99.1%; Pred. No. 1.5e-141;
Matches 779; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 920 tggatatttgaagaaaaaaatcctcaaggaagcctaaataatgattgctacacaaag 979
    |||||
DB 26917 TTGTTTTCAGAAAAAAATCTCAGAGAAAAAGCAATTAAGATTCCTACACAAAG 26858
    |||||

QY 980 tgacaaatttaagaaagccttcagctgattggtggaattcagaagaaaaata 1039
    |||||
DB 26857 TGACCAATTTTAAAGAGCCTTCATGAGCTGATGCTGGGAAATTCAGAAAAAAATA 26798
    |||||

QY 1040 caggaagaacacacacagagaaggttttctcctacacccgcaagaacatataga 1099
    |||||
DB 26797 CAGGAAAGAACACACAGAGAGGTTTTTCTTCAACCGCAAGCAATATATTAGA 26738
    |||||

QY 1100 tacatgaatcctaatatattatgagcaatttgcatttatttcaaatatcattgt 1159
    |||||
DB 26737 TACATGAATCTCAATTAATATATATGCAATTAATTCATTTTCAAAATTAATTGT 26678
    |||||

QY 1160 gggagacatgaatcctcttgagcaatcgtatatttgggaagccttcaaaagtacaa 1219
    |||||
DB 26677 GGGGACATGATATCTCTTGAGCAATCTGATATTTTGGGAAGTCTTTAAAAAGTTACAA 26618
    |||||

QY 1220 attatcaataatattactaagatagaatgaatcagaacaaagaacacagaatt 1279
    |||||
DB 26617 ATTATCAATAATTAATTACTAGTAGATGAATGATTCAGAAAAAAGAAATCAGCAAT 26558
    |||||

QY 1280 aggtgtggtgctggtggtgtaagaacacatgagcaatttcaaatgcaaaagtc 1339
    |||||
DB 26557 AGGATGTGGCTGGCTGCTGATGAAGCAGCATGATGAAATTCATTAAGTTGCAAAAGTC 26498
    |||||

QY 1340 aaacaatctgacatgacacacagaaatcaaaaataatcagaataatagagactatata 1399
    |||||
DB 26497 AAAACATATCTGTAACATGACAGCAAGAAATCAAAATTAATTCAGAAATAGAGACTTATA 26438
    |||||

QY 1400 aatgcatatataatgatacttgcataataaagccattggaagaaaggaagattaga 1459
    |||||
DB 26437 AATGCAATTAATACATGATACCTTTTACATTAATTAAGCCATTTGGAAGAAACGGAAGATTAGA 26378
    |||||

QY 1460 taactaataactgactctctcttgcataaactgataatgataatgattgattactt 1519
    |||||
DB 26377 TACTAATAATACATTTGACTATCTCTTTGTAATAAGAGTACTAAATGATGTTTACTT 26318
    |||||

```

```

QY 1520 tccatggtggaatttaattactttctcttgtaattttctcctgatatattaaaca 1579
    |||||
DB 26317 TCCATGTGGGAATTTTAATTAATCTTTCTTTGTAATTTTCTCTGTAATTTTAAACA 26258
    |||||

QY 1580 aatactggtatagtttacatatataagaatattgctcaattgaaaggcaaggcca 1639
    |||||
DB 26257 AATAGCTGATAGTTTACATATTTTAAAGATTTGTTCAATTAAGAGCAAGGCCA 26198
    |||||

QY 1640 ggtcagcaattttcaactgatagtacattaaataaataactaataaataaata 1699
    |||||
DB 26197 GGTTCAGCAATTTTCAAACTGATGATTAATTAATAATTAATTAATAATTA 26138
    |||||

QY 1700 tatttc 1705
    |||||
DB 26137 TATTTTC 26132
    |||||

```

```

RESULT 3
AP001880      169776 bp      DNA      linear      PRI 26-APR-2001
LOCUS
DEFINITION      Homo sapiens genomic DNA, chromosome 11q, clone:CTD-3110P2,
ACCESSION      AP001880
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Homo sapiens DNA, clone:CTD-3110P2.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (sites)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens genomic DNA
Published Only in Database (2000) In press
2 (bases 1 to 169776)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (24-APR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho,Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gs.c.riken.go.jp, URL:http://hnp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Apr 26, 2001 this sequence version replaced gi:10716820.
FEATURES
source
Location/Qualifiers
1..169776
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="CTD-3110P2"

```

```

BASE COUNT      48281 a 32818 c 33507 g 55170 t
ORIGIN

Query Match      45.4%; Score 774.8; DB 9; Length 169776;
Best Local Similarity 99.1%; Pred. No. 1.4e-141;
Matches 779; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 920 tggatatttgaagaaaaaaatcctcaaggaagcctaaataatgattgctacacaaag 979
    |||||
DB 82466 TTGTTTTCAGAAAAAAATCTCAGAGAAAAAGCAATTAAGATTCCTACACAAAG 82525
    |||||

QY 980 tgacaaatttaagaaagccttcagctgattggtggaattcagaagaaaaata 1039
    |||||
DB 82526 TGACCAATTTTAAAGAGCCTTCATGAGCTGATGCTGGGAAATTCAGAAAAAAATA 82585
    |||||

QY 1040 caggaagaacacacacagagaaggttttctcctacacacagcaagaacatataga 1099
    |||||
DB 82586 CAGGAAAGAACACACAGAGAGGTTTTTCTTCAACCGCAAGCAATATATTAGA 82645
    |||||

QY 1100 tacatgaatcctaatatattatgcatatttgcatttatttcaaatatcattgt 1159
    |||||

```



```

misc_feature      /note="assembly-fragment"
                  140741..152854
misc_feature      /note="assembly-fragment"
                  152855..161965
misc_feature      /note="assembly-fragment clone_end:17 vector_side:left"
                  162066..165410
misc_feature      /note="assembly-fragment clone_end:16 vector_side:left"
                  165511..166848
BASE COUNT      55669 a 32391 c 31923 g 46064 t 801 others
ORIGIN

```

```

Query Match      44.7%; Score 762.8; DB 2; Length 166848;
Best Local Similarity 99.0%; Pred. No. 3.2e-139;
Matches 778; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 920 tggattttggaagaaatcgaagaaagcgaagaaatgattgctacacaaag 979
Db 36630 TGTGTTTTCAGAAAAAAATCTCAAGAAAGCATATAAAATGATGCTACACAAAG 36571
QY 980 tggcaaatcttaagaagccttcagctgattggtggaagaaatcagaagaaataa 1039
Db 36570 TGACCAAAATTTTAAAGAGCCCTTCATGAGCTGATGGTGGGAAATTCAGAAAAAAATA 36511
QY 1040 caagaaagaacacacacagaagaggtttttccctacacacagcaagacatatatta 1099
Db 36510 CAGGAAAAAGAACACACAGAGGGTTTTTTCCTACACAGCAAGAAATATATAGA 36451
QY 1100 taatgaatcgaatataatattatgacattatgacattatcctcaaatattactgt 1159
Db 36450 TACATGATATCTCAATATATATATGTCATTTATTTGCAATTTATTAATCTGT 36391
QY 1160 ggggaatgtaactctctgaagcaatcgtatcttgggaagtcctcctaaagttaca 1219
Db 36390 GGGGACAGTATCTCTGAGCAGATCTGATTTTGGGAAGTCCTTTAAAAAGTTTACA 36331
QY 1220 attatcaataaattacagagagatgaatgattcgaagaaacaaagaaatcacagaat 1279
Db 36330 ATTTATCAATTAATTTACTAGAGATGATGATGATTCAGAAACAAAGAAATACAGAAAT 1280
QY 1280 aggatgtggtcgtgctggtgatgaagacacatgattgattcacaatgagttacatt 1339
Db 36270 AGGATGTGGCTGCTGCTGATGATGACACCATGCTGATGATGATGATGATGATGATGAT 1340
QY 1340 aaaaactactgacatgacacacagaacaaataaataccagaatagagactatata 1399
Db 36210 AAAACATATCTGATGATGACACCAAGAAATCAAAATTAATCCAGAAATAGACCTATATA 1400
QY 1400 aatgcaatataacatgactctttgacataataagcattggaagaaagcgaataga 1459
Db 36150 AATGCAATTAATACATGATGATCTTTGACATTAATAGCCATTGGAAAAAGGAAATAGA 1460
QY 1460 tacttaataaactgactatctcttgaataagtcactaaatgattgattacttt 1519
Db 36090 TACTTAATAAATGATGATGATCTTTGATTAATAGCTCACTAAATGATGATGATGATGAT 1530
QY 1530 tcaatggtgaatattactactcttcttcttcttcttcttcttcttcttcttctt 1579
Db 36030 TCCATGCTGGAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1580
QY 1580 aatgagcgtgataagttacataatataaagatatgcttgaatggaagcgaagccca 1639
Db 35971 AATAGCTGATATAGTTTCAATATATATATAATTTGTTCAAAATTTGAAGGCAAGGCCA 1640
QY 1640 gttcagaatcttcaacgtatgacatttaataaataaataaataaataaataaataa 1699
Db 35911 GGTTCAGCAATTTCAAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1700
QY 1700 tatttc 1705
Db 35851 TATTTT 35846

```

```

RESULT 5
AC073256/c
LOCUS
DEFINITION
AC073256
Home sapiens chromosome 11 clone Repl-204K20, WORKING DRAFT
AC073256
SEQUENCE, 20 unordered pieces.
AC073256
AC073256.2 GI:8569847
HTG; HTGS_PHASE1; HTGS_DRAFT.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 170682)
The sequence of Homo sapiens clone
2 (bases 1 to 170682)
Waterston, R.H.
Direct Submission
Submitted (12-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108 USA
On Jun 16, 2000 this sequence version replaced gi:8468975.

```

COMMENT

```

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0204K20
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye primer ET; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 157404 bases at least Q40
Consensus quality: 160944 bases at least Q40
Consensus quality: 163054 bases at least Q30
Insert size: 161000; agarose-fp
Insert size: 168782; sum-of-ctligs
Quality coverage: 4.62 in Q20 bases; agarose-fp
Quality coverage: 4.50 in Q20 bases; sum-of-ctligs
----- NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2201: contig of 2200 bp in length
2300: gap of unknown length
3481: contig of 1181 bp in length
3582: gap of unknown length
4832: contig of 1251 bp in length
4932: gap of unknown length
4933: contig of 1698 bp in length
6631: gap of unknown length
6731: contig of 1698 bp in length
11545: gap of unknown length
11645: contig of 4815 bp in length
11646: gap of unknown length
11646: contig of 4171 bp in length
15817: gap of unknown length
15917: contig of 3982 bp in length
19898: gap of unknown length
19998: contig of 4952 bp in length
24950: gap of unknown length
24951: contig of 5299 bp in length
30349: gap of unknown length
30350: contig of 6012 bp in length
30450: gap of unknown length
36462: gap of unknown length

```

1

1

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL

Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Homo sapiens 116,273 genomic DNA of 11q13
2 (bases 1 to 116273)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (12-OCT-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagami-cho, Kanagawa 228-8555,
Japan (E-mail: hattori@gs.c.riken.go.jp,
URL: http://hgp.gs.c.riken.go.jp/,
Fax: 81-42-778-9924)
On May 31, 2000 this sequence version replaced gi:6997486.

COMMENT

Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gs.c.riken.go.jp/
Contact: hattori@gs.c.riken.go.jp
Center project name: HumDrafl1
Center clone name: Xpac-662A14
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 103806 bases at least Q40
Consensus quality: 109449 bases at least Q30
Insert size: 1134/73; sum-of-coverage
Quality coverage: 4.03x in Q20 bases; sum-of-coverage

NOTE: This is a 'working draft' sequence. It currently consists of
29 contigs. The true order of the pieces is not known and the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved

```

1 38239 38138 contig of 38138 bp in length
44799 44698 contig of 6460 bp in length
49615 49514 contig of 4716 bp in length
55635 55534 contig of 5920 bp in length
60237 60136 contig of 4502 bp in length
65724 65623 contig of 5387 bp in length
70203 70102 contig of 4379 bp in length
71488 71387 contig of 1185 bp in length
73940 73839 contig of 2352 bp in length
77467 77366 contig of 3427 bp in length
80166 80065 contig of 2599 bp in length
83926 83825 contig of 3660 bp in length
86878 86777 contig of 2852 bp in length
89884 89783 contig of 2906 bp in length
92852 92751 contig of 2668 bp in length
95169 95068 contig of 2217 bp in length
96740 96739 contig of 1571 bp in length
98445 98344 contig of 1732 bp in length
99845 99744 contig of 1073 bp in length
101254 101153 contig of 1309 bp in length
102868 102767 contig of 1615 bp in length
104921 104820 contig of 1953 bp in length
106500 106400 contig of 1479 bp in length
108529 108428 contig of 1929 bp in length
110941 110840 contig of 2212 bp in length
112467 112366 contig of 1426 bp in length
113765 113664 contig of 1198 bp in length
114974 114873 contig of 1209 bp in length
115075 114974 contig of 1209 bp in length
116273 116273 contig of 1199 bp in length

```

Sequence updated (26-May-2000).

* NOTE: This is a 'working draft' sequence. It currently
* consists of 29 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

FEATURES

```

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 38138: contig of 38138 bp in length
38239 38238: gap of 100 bp
44698 44697: contig of 6460 bp in length
44799 44798: gap of 100 bp
49514 49513: contig of 4716 bp in length
49615 49614: gap of 100 bp
55534 55533: contig of 5920 bp in length
55635 55634: gap of 100 bp
60136 60135: contig of 4502 bp in length
60237 60236: gap of 100 bp
65623 65622: contig of 5387 bp in length
65724 65723: gap of 100 bp
70102 70101: contig of 4379 bp in length
70203 70202: gap of 100 bp
71387 71386: contig of 1185 bp in length
71488 71487: gap of 100 bp
73839 73838: contig of 2352 bp in length
73940 73939: gap of 100 bp
77366 77365: contig of 3427 bp in length
77467 77466: gap of 100 bp
80065 80064: gap of 100 bp
80166 80165: gap of 100 bp
83825 83824: contig of 3660 bp in length
83926 83925: gap of 100 bp
86777 86776: contig of 2852 bp in length
86878 86877: gap of 100 bp
89783 89782: contig of 2906 bp in length
89884 89883: gap of 100 bp
92751 92750: contig of 2668 bp in length
92852 92851: gap of 100 bp
95068 95067: contig of 2217 bp in length
95169 95168: gap of 100 bp
96739 96738: contig of 1571 bp in length
96740 96739: gap of 100 bp
97912 97911: contig of 1073 bp in length
97913 97912: gap of 100 bp
98013 98012: contig of 1732 bp in length
98445 98444: gap of 100 bp
99845 99844: contig of 1309 bp in length
101153 101152: gap of 100 bp
101254 101253: contig of 1615 bp in length
102868 102867: gap of 100 bp
102969 102968: contig of 1953 bp in length
104921 104920: gap of 100 bp
105022 105021: contig of 1479 bp in length
106500 106499: gap of 100 bp
106501 106500: contig of 1929 bp in length
108529 108528: gap of 100 bp
108530 108529: contig of 2212 bp in length
110841 110840: gap of 100 bp
110942 110941: contig of 1426 bp in length
112367 112366: gap of 100 bp
112468 112467: contig of 1198 bp in length
113665 113664: gap of 100 bp
113766 113765: contig of 1209 bp in length
114975 114974: gap of 100 bp
115075 115074: gap of 100 bp
116273 116272: contig of 1199 bp in length.
Location/Qualifiers
1..116273
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
/clone="Xpac-662A14"
1..38138
/note="assembly-fragment"
38239..44698
/note="assembly-fragment"
44799..49514

```

misc_feature
misc_feature
misc_feature
misc_feature

[illegible]

```

RESULT 7
AP003420
LOCUS
DEFINITION
Homo sapiens genomic DNA, chromosome 11q, clone:RP11-614K12,
complete sequence.
ACCESSION
AP003420
VERSION
AP003420.2 GI:15320509
KEYWORDS
HTG.
SOURCE
Homo sapiens DNA, clone:RP11-614K12.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 160945)
AUTHORS
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Matsunabe,H. and Sakaki,Y.
Homo sapiens genomic DNA
Published Only in Database (2001) In press
2 (bases 1 to 160945)
AUTHORS
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Matsunabe,H. and Sakaki,Y.
Direct Submission
TITLE
Submitted (16-MAR-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan
1-7-22 Suenhiro-chou,Tsuriumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT
On Aug 27, 2001 this sequence version replaced gi:13383333.
FEATURES
source
1..160945
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-614K12"
BASE COUNT
46743 a 31738 c 32047 g 50417 t
ORIGIN

```

Query Match	34.5%	Score 588,	DB 9;	Length 160945;
Best Local Similarity	100.0%;	Pred. No. 3.8e-105;		
Matches 588;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0


```

misc-feature      152935..161965
                  /note="assembly_fragment clone_end:17 vector_side:left"
misc-feature      162066..165410
                  /note="assembly_fragment clone_end:56 vector_side:left"
misc-feature      165511..166848
                  /note="assembly_fragment"
BASE COUNT      55669 a 32391 c 31923 g 46064 t      801 others
ORIGIN

```

```

Query Match      34.5%; Score 588; DB 2; Length 166848;
Best Local Similarity 100.0%; Pred. No. 3.8e-105; Indels 0; Gaps 0;
Matches 588; Conservative 0; Mismatches 0;

```

```

OY 1 aagaactgaggaagccttccactacgctatgacatgagtcgagccacgag 60
    |||||
Db 161119 AAGAACTGAGAGGCTCTTCCACTACGCGCTGATGCACTGGTGTGCGGCCCATAG 161178
    |||||
OY 61 atggaactgatgcgagagcaatacaagcttaattgaataaagaagaatttt 120
    |||||
Db 161179 ATGGAATGATGCGAGATCAATACAGCTTAATTGATTAATAAGAAATATT 161238
    |||||
OY 121 tctcccttgacatctctcgttaagacatgtgctcctctctggggggtcagcgtgtca 180
    |||||
Db 161239 TCTCCCTTGAACTTATCTCCGTAAACCATGTGGCTCTTGCGGGGTCAAGCTTCA 161298
    |||||
OY 181 caatcaatgaccttgagagagctcttgatcaagttgagagccttgaggattcagatg 240
    |||||
Db 161299 CAATCAATGAGGCTTGAAGAGACCTTGATGATCAAGTTGAGAGGCTTGAGAAATTTACAGATC 161358
    |||||
OY 241 ctccatcgtgtttatctctcctctcctcgtatcaatccctcatalactatgtagag 300
    |||||
Db 161359 CTTCATCTGGTTTATCTCTCCCTCTCTCATGTTAATCCCTCATATACCTGCTAGAG 161418
    |||||
OY 301 aacttgctgcagccatctcctgctgctatgctgtgctgctcagatgcagcaataact 360
    |||||
Db 161419 AACTTTGCTGCGACCATCTCTGCTCATGTTGCTGGTCCACATGCTGGACATTAATTAATCT 161478
    |||||
OY 361 ggaactggtgaatgaactggaactcctcgtgtaagaatgccccttgagaatcctatccca 420
    |||||
Db 161479 GGAATCTGGAATGAACCTGGAATCTCACTGAAGATGCCCTTTGAGAACTCTATCCCA 161538
    |||||
OY 421 ctgactcaaatctgcagcagagagatgctgctgcttgcattccatcccaatgagagctt 480
    |||||
Db 161539 CTGCACTCAAAATGTGAGCGACAGAGAGTGTGCTGCTTGTCCATCCCACTGCGAGCTT 161598
    |||||
OY 481 ctccactgaaatggagactatccacagacaaagtggagagacagaaacctgtgtgagat 540
    |||||
Db 161599 CTTCACTGAATGGAGCTATCCACAGACAAAGTGAAGGACAGACAGAACCCGTGTGGAT 161658
    |||||
OY 541 ggcctggatataatgataaagctactcctcctgacacattgtgactaag 588
    |||||
Db 161659 GCGCTGGTATATGATCAAAAGCTTCCCTTGACCATTTGTGACTAAG 161706
    |||||

```

```

RESULT 9
AP000596      116273 bp      DNA      linear      HMG 30-MAY-2000
LOCUS      Homo sapiens chromosome 11 clone XHPac-662A14 map 11q13, WORKING
DEFINITION      DRAFT SEQUENCE, 29 unordered pieces.
ACCESSION      AP000596
VERSION      AP000596.2 GI:8118802
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 116273)
AUTHORS      Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      Homo sapiens 11q13 genomic DNA of 11q13
JOURNAL      Published Only in Database (1999) In press
REFERENCE      2 (bases 1 to 116273)

```

AUTHORS

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE

Submitted (12-0CT-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@gsc.riken.go.jp, Tel:81-42-778-9923,
URL:http://hgp.gsc.riken.go.jp/, Fax:81-42-778-9924)

COMMENT

On May 31, 2000 this sequence version replaced gl:6597486.
----- Genome Center

Center: RIKEN Genomic Sciences Center(GSC)

Center code: RIKEN

Web site: http://hgp.gsc.riken.go.jp/

Contact: hattori@gsc.riken.go.jp

----- Project Information

Center project name: HumDrafl1

Center clone name: XHPac-662A14

----- Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 103806 bases at least Q40

Consensus quality: 109449 bases at least Q30

Consensus quality: 111869 bases at least Q20

Insert size: 113473; sum-of-ctrls

Quality coverage: 4.03x in Q20 bases; sum-of-ctrls

NOTE: This is a 'working draft' sequence. It currently consists of

29 contigs. The true order of the pieces is not known and their

order in this sequence record is arbitrary. Gaps between the

contigs are represented as runs 'N', but the exact sizes of the gaps

are unknown. This record will be updated with the finished sequence

as soon as it is available and the accession number will be

preserved

1 38138 contig of 38138 bp in length

38239 44689 contig of 44689 bp in length

44799 4716 bp in length

49615 55534 contig of 55534 bp in length

55635 60136 contig of 60136 bp in length

60237 65623 contig of 65623 bp in length

65724 70102 contig of 70102 bp in length

70203 71387 contig of 71387 bp in length

71488 73839 contig of 73839 bp in length

73940 77366 contig of 77366 bp in length

77467 80065 contig of 80065 bp in length

80166 83825 contig of 83825 bp in length

83926 86777 contig of 86777 bp in length

86878 89751 contig of 89751 bp in length

89884 92852 contig of 92852 bp in length

92852 95068 contig of 95068 bp in length

95169 96739 contig of 96739 bp in length

96840 97912 contig of 97912 bp in length

98013 99744 contig of 99744 bp in length

99845 101153 contig of 101153 bp in length

101254 102868 contig of 102868 bp in length

102969 104521 contig of 104521 bp in length

103022 106500 contig of 106500 bp in length

108329 108529 contig of 108529 bp in length

108601 110841 contig of 110841 bp in length

110942 112367 contig of 112367 bp in length

112468 113665 contig of 113665 bp in length

113766 114974 contig of 114974 bp in length

115075 116273 contig of 116273 bp in length

Sequence updated (26-May-2000).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 29 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of 'N', but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

*	1	38138: contig of 38138 bp in length
*	38139	38238: gap of 100 bp
*	38239	44698: contig of 6460 bp in length
*	44699	44798: gap of 100 bp
*	44799	49514: contig of 4716 bp in length
*	49515	49614: gap of 100 bp
*	49615	55534: contig of 5920 bp in length
*	55535	55634: gap of 100 bp
*	55635	60136: contig of 4502 bp in length
*	60137	60236: gap of 100 bp
*	60237	65623: contig of 5587 bp in length
*	65624	65723: gap of 100 bp
*	65724	70102: contig of 4379 bp in length
*	70103	70202: gap of 100 bp
*	70203	71387: contig of 1185 bp in length
*	71388	71487: gap of 100 bp
*	71488	73839: contig of 2352 bp in length
*	73840	73939: gap of 100 bp
*	73940	77366: contig of 3427 bp in length
*	77367	77466: gap of 100 bp
*	77467	80065: contig of 2599 bp in length
*	80066	80165: gap of 100 bp
*	80166	83825: contig of 3660 bp in length
*	83826	83923: gap of 100 bp
*	83926	86777: contig of 2852 bp in length
*	86778	86877: gap of 100 bp
*	86878	89783: contig of 2906 bp in length
*	89784	89883: gap of 100 bp
*	89884	92751: contig of 2866 bp in length
*	92752	92851: gap of 100 bp
*	92852	95068: contig of 2217 bp in length
*	95069	95168: gap of 100 bp
*	95169	96739: contig of 1571 bp in length
*	96740	96839: gap of 100 bp
*	96840	97912: contig of 1073 bp in length
*	97913	98012: gap of 100 bp
*	98013	99744: contig of 1732 bp in length
*	99745	99844: gap of 100 bp
*	99845	101153: contig of 1109 bp in length
*	101154	101253: gap of 100 bp
*	101254	102868: contig of 1615 bp in length
*	102869	102968: gap of 100 bp
*	102969	104921: contig of 1953 bp in length
*	104932	105021: gap of 100 bp
*	105022	106500: contig of 1479 bp in length
*	106501	106600: gap of 100 bp
*	106601	108529: contig of 1929 bp in length
*	108530	108629: gap of 100 bp
*	108630	110841: contig of 2212 bp in length
*	110842	110941: gap of 100 bp
*	110942	112367: contig of 1426 bp in length
*	112368	112467: gap of 100 bp
*	112468	113665: contig of 1198 bp in length
*	113666	113765: gap of 100 bp
*	113766	114974: contig of 1209 bp in length
*	114975	115074: gap of 100 bp
*	115075	116273: contig of 1199 bp in length.
Location/Overlap/Length		

```
misc_feature /note="assembly-fragment" 60237. .65623
misc_feature /note="assembly-fragment" 65274. .70102
misc_feature /note="assembly-fragment" 70203. .71387
misc_feature /note="assembly-fragment" 71488. .73839
misc_feature /note="assembly-fragment" 73940. .77366
misc_feature /note="assembly-fragment" 77467. .80065
misc_feature /note="assembly-fragment" 80166. .83825
misc_feature /note="assembly-fragment" 83926. .86777
misc_feature /note="assembly-fragment" 86878. .89783
misc_feature /note="assembly-fragment" 89884. .92751
misc_feature /note="assembly-fragment" 92852. .95068
misc_feature /note="assembly-fragment" 95169. .96739
misc_feature /note="assembly-fragment" 96840. .97912
misc_feature /note="assembly-fragment" 98013. .99744
misc_feature /note="assembly-fragment" 99845. .101153
misc_feature /note="assembly-fragment" 101254. 1.02868
misc_feature /note="assembly-fragment" 102969. 1.04921
misc_feature /note="assembly-fragment" 105022. 1.06500
misc_feature /note="assembly-fragment" 106601. 1.08529
```

Query Match	34.48	Score 586.4	DB 2	Length 116273
Best Local Similarity	99.88	Pred. No. 8.2e-105		
Matches 587	Conservative	0	Mismatches 1	Gaps 0
QY	1	aagactgagaaactcttccactacgctgtatgcaactggttgatgcgcggccatgg	60	
Db	111636	AAGAACTGAGGAACCTCTTCCACTACGCTGTATGACATGGGTGAGTCCGGCCCATATG	111695	
QY	61	atgagaattgagcgcgagatcaatacaagaacttaattgaattaataaagaagaattt	120	
Db	111696	ATGGAATATGATGCGGAGATCAATACAAAGCTTAATTGTGAATTAATAAAGAAATATTT	111755	
QY	121	tctcccttgaactatctccgfaaagccatgtgctccctctctggggatcagctgtca	180	
Db	111756	TCTCCCTTGAACCTATATCTCCGTAAGCAATGTGCTCTCTTGGGGGTCCAGTGTCA	111815	
QY	181	caatcaatgaccttgaagagcctctgaagtcgaagtggagccttgggaagattcaagt	240	
Db	111816	CAATCAATAGGCTTTATGAGAGCTCTGTGATCAAGTTGAGAGCCTTGGAGATTTCCAGATG	111875	
QY	241	cttcaactggttttattcttccctctctcatalgtatttaaccctcatatactgtag	300	
Db	111876	CTTCAATCTGTTTTATTTCTTCTTCCCTCTCCATGTTATTAATTCCTCAATATCTCTAGAG	111935	
QY	301	aacttgtcgcagccatctccgtgatcgtttgcgggggtccacatgctcgacaataact	360	
Db	111936	AACCTTGTGCGACCATTCCTGCTGATCTGTGCTGGGGTCCACATGCGGCAATAAATACT	111995	
QY	361	ggatctggttaatgaactggaactcctcaatgaaagatgccctcttgaaatcttatocca	420	
Db	111996	GGATCTGGTATGAACATGGAATCCTCAGTGAAGATCCCTCTTGAGAAATCTTATGCCA	112055	
QY	421	ctagactcaaatctgagccagagaagtgcgtgcctttgcatatcccaatgtagactt	480	

RESULT 13
AX074149 1977 bp DNA linear PAT 06-FEB-2001
LOCUS AX074149
DEFINITION Sequence 5 from Patent WO0104283.
AX074149
VERSION AX074149.1 GI:12710361
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1977)
AUTHORS Sun, W.
TITLE Organic anion transporter genes and proteins
JOURNAL Patent: WO 0104283-A 5 18-JAN-2001;
Metabasis Therapeutics Inc. (US)
FEATURES
source 1..1977
Location/Qualifiers
BASE COUNT 502 a 491 c 450 g 534 t
ORIGIN
Query Match 22.2%; Score 378; DB 6; Length 1977;
Best Local Similarity 75.9%; Pred. No. 5.9e-64;
Matches 481; Conservative 0; Mismatches 150; Indels 3; Gaps 1;
QY 64 agaaatgatgcgaggaatcaatacaagctaatgtgaataaagaataatttct 123
DB 113 AGAGAGTGGGGTGGAGATCAAAACATTTAGTGTAGCTTAGGAAAGAAACATTTTCC 172
QY 124 ccccttgaaactatctccgttaagcaatgtgcctcctcttgagggtcaagtgtaaca 183
DB 173 CTTTGTGAACCTCTGTGATACGATTTTCTCTACTTGAAGATCACTGTCAAC 232
QY 184 tcaatgagccttgagagctcttgagcaagttgagagccttgagagattcagatgctt 243
DB 233 TCAATGGCTTTCAAGAGCTCTCTGGGTCAAGCTGTGACCTTGGAGATTCAGATCTT 292
QY 244 catctggtttatctctcctcctcctcctcctcctcctcctcctcctcctcctcct 303
DB 293 CAGACTGTTTCTCTCAATCTTGTGCTGATACATACCTTCAATTTTGTGGAGAAC 352
QY 304 ttgtgtgagcattccctcctcctcctcctcctcctcctcctcctcctcctcctcct 363
DB 333 TTCACTGATTCATACCTGGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 412
QY 364 tctggttaatgaactgaatcctcctcctcctcctcctcctcctcctcctcctcctcct 423
DB 413 TCTGACATGACACTGGGGCCCTCAGCAAGATGCACTTTGAGATCTCCACTG 472
QY 424 gactcaaatctgagagcagagaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 483
DB 473 GACTCAAAACATGAGGCCAGAGAGAGTGTGCTGCTTGTTCATCTCAGAGCAGCTCT 532
QY 484 caccgtgaatgagactcaccagcaagttgagagagcagagaaacctgtgtgtgtgtgt 543
DB 533 CACCTGAATGGAGCTTCCCAACACAGAGTGAAGCAGACATGAGACCTCTGTGTGATG 592
QY 544 tgggtatataatgaactgaactcctcctcctcctcctcctcctcctcctcctcctcct 600
DB 593 TGGGTGATGACAGAACTCTCTTCATCCACACATCGGTGAAGTGGATGTGTA 652
QY 601 tttgtatataatgaactgaactcctcctcctcctcctcctcctcctcctcctcctcct 660
DB 653 TGTGACTCTCATCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 712
QY 661 ggaagcaccatgaatgt 694
DB 713 GGAGGACATCTTAGCGGTTCATTTATCAGACAGGT 746

RESULT 14
AP003420/c 160945 bp DNA linear PRI 28-AUG-2001
LOCUS AP003420/c
DEFINITION Homo sapiens genomic DNA, chromosome 11q, clone:RP11-614K12,
complete sequence.
AP003420
VERSION AP003420.2 GI:15320509
KEYWORDS HTG.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 160945)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Published Only in Database (2001) In press
JOURNAL 2 (bases 1 to 160945)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
REFERENCE Direct Submission
TITLE Submitted (16-MAR-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suenho-cho, Tsukuba, Ibaraki, Japan, 305-8565, Japan
(E-mail: hattori@gs.c.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT On Aug 27, 2001 this sequence version replaced gi:13383333.
FEATURES
source 1..160945
Location/Qualifiers
BASE COUNT 46743 a 31738 c 32047 g 50417 t
ORIGIN
Query Match 20.0%; Score 340.8; DB 9; Length 160945;
Best Local Similarity 76.6%; Pred. No. 6e-57;
Matches 444; Conservative 0; Mismatches 132; Indels 4; Gaps 2;
QY 11 gaagctcttccactcagcgtgtatgtgacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 70
DB 83641 GAGGAATTTGCTTCCCTGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 83582
QY 71 gatgcagagatca--atacaagcttaattgaattgaattgaattgaattgaattgaatt 128
DB 83581 GGGGTCAGGATCAACACACACACTTGTGTGATTTAAGGAAGAAATTTTCT--CTT 83524
QY 129 tgaactatctccgttaagcaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 188
DB 83523 TGAACCTTATCTGGATACACTGATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 83464
QY 189 ggccttttgagagcctctgtgagcagttgagagcctctgtgagattcgaatgtcattct 248
DB 83463 GGCCTTTCAAGACCTCTCAATGATGATGATGATGATGATGATGATGATGATGAT 83404
QY 249 gtttttattctctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 308
DB 83403 GGTTCCTTATTAAGTTGTCAGGTGATTAATTAATTAATTAATTAATTAATTAATTA 83344
QY 309 tggagcattcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 368
DB 83343 ACCATTTCATACCTTGAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 83284
QY 369 taatgaactgaatccatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 428
DB 83283 CATATACCTTGGAGCCCTCAAGCAGAGATCCTCTGCAAAATCTCATCTCCATTGACTC 83224
QY 429 aacttgagccagagaagt 488

Db 83223 AAATCTGAGCCAGAGAGTGTCTGCTTTGCTTCATCCCAAGTGAAGCTCATTCATCT 83164
 Oy 489 gaatggagctctccacagacaaatgagcagacaaacccctgtgtgatgctgggt 548
 Db 83163 GAATGGAGCTTCCCAACACGATGAGCCAGTACAGAGCCCTGTGTGATGCTGGGT 83104
 Oy 549 atatcatcaagctactctccctgcacattgtgactaa 588
 Db 83103 AATATGACCAAGCTCTTCCCTTCACCATTTGAGACTAAG 83064
 RESULT 15
 AC090778 164809 bp DNA linear HTG 27-JAN-2002
 LOCUS Homo sapiens chromosome 11 clone CTD-3110P2 map 11, WORKING DRAFT
 DEFINITION AC090778.2 GI:14336491
 ACCESSION AC090778.2
 VERSION AC090778.2
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 164809)
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 TITLE Homo sapiens chromosome 11, clone CTD-3110P2
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 164809)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
 Batra, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A.,
 Camarata, J., Campopiano, A., Choquel, Y., Colangelo, M., Collins, S.,
 Collamore, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J.S.,
 Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
 Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Karatas, A., Larocque, K., Lamazares, R., Landers, T.,
 Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P.,
 Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
 McSheeters, R., Meldrum, J., Menes, L., Mihova, T., Mlenga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
 Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roselli, M.,
 Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,
 Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strausz, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Travers, M., Travis, N., Trifoglio, J., Vassiliev, H., Viel, R., Vo, A.,
 Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
 Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (10-MAR-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jun 11, 2001 this sequence version replaced gi:13270647.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RW/RepeatMasker.html
 Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L12735
 Center clone name: 3110.P.2
 Summary Statistics
 Sequencing Vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly Program: Phrap; version 0.960731
 Consensus quality: 157859 bases at least Q40
 Consensus quality: 161528 bases at least Q30
 Consensus quality: 162934 bases at least Q20
 Insert size: 172000; agarose-IP
 Insert size: 163409; sum-of-contigs

Quality coverage: 5.0 in Q20 bases; agarose-IP
 Quality coverage: 5.2 in Q20 bases; sum-of-contigs

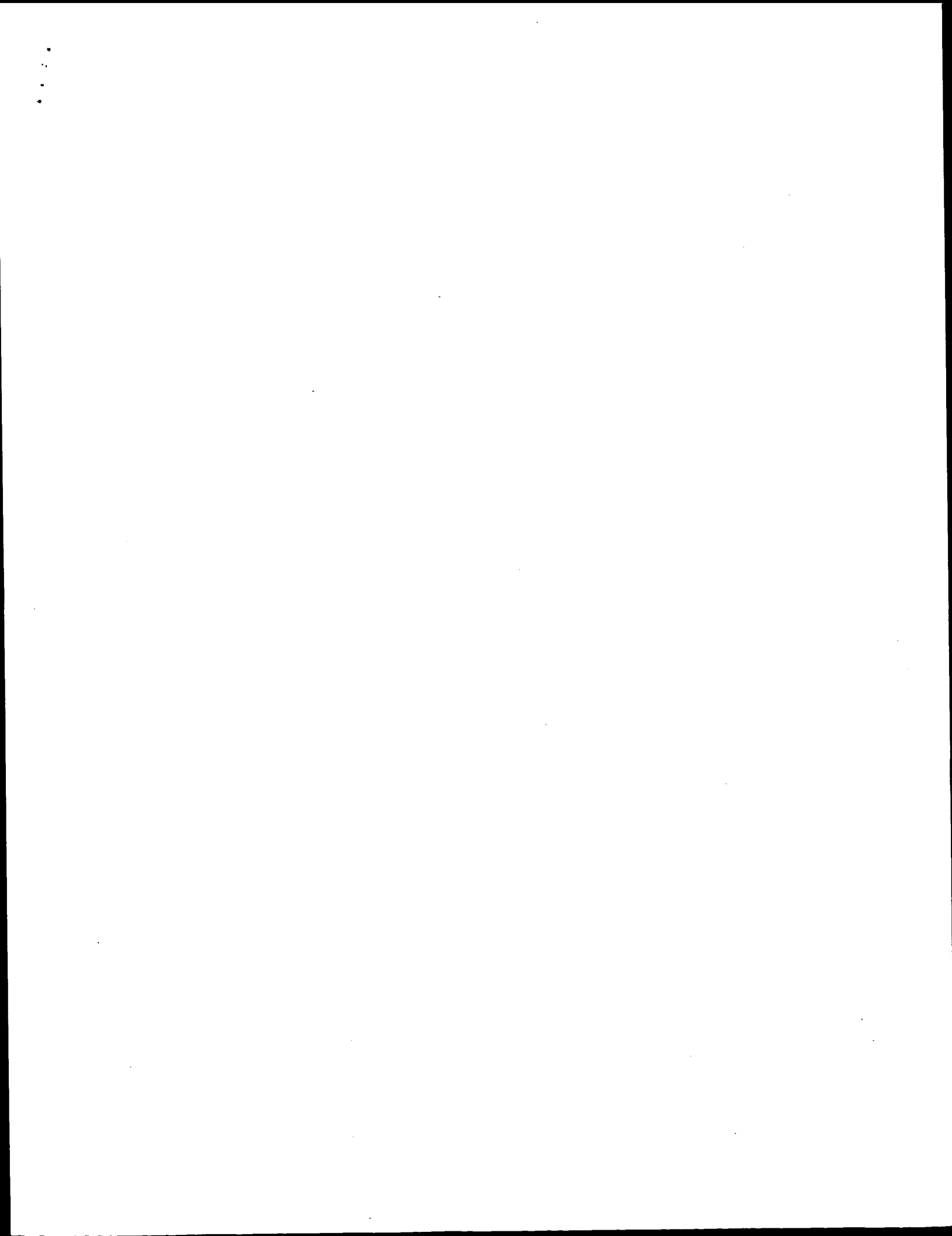
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 15 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 1 1069: contig of 1069 bp in length
 1070 1169: gap of 100 bp
 1170 2379: contig of 1210 bp in length
 2380 2479: gap of 100 bp
 2480 3440: contig of 961 bp in length
 3441 3540: gap of 100 bp
 3541 4880: contig of 1340 bp in length
 4881 4980: gap of 100 bp
 4981 6768: contig of 1788 bp in length
 6769 6868: gap of 100 bp
 6869 8318: contig of 1450 bp in length
 8319 8418: gap of 100 bp
 8419 10046: contig of 1628 bp in length
 10047 10146: gap of 100 bp
 10147 12455: contig of 2309 bp in length
 12456 12555: gap of 100 bp
 12556 15164: contig of 2609 bp in length
 15165 15264: gap of 100 bp
 15265 16648: contig of 1384 bp in length
 16649 16748: gap of 100 bp
 16749 29429: contig of 12681 bp in length
 29430 29529: gap of 100 bp
 29530 41226: contig of 11697 bp in length
 41227 41326: gap of 100 bp
 41327 57370: contig of 16044 bp in length
 57371 57470: gap of 100 bp
 57471 102968: contig of 45498 bp in length
 102969 103068: gap of 100 bp
 103069 164809: contig of 61741 bp in length.
 FEATURES
 source
 1..164809
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="CTD-3110P2"
 /clone_lib="CITD Human BAC"
 1..1069
 /note="assembly-fragment"
 1170..2379
 /note="assembly-fragment"
 2480..3440
 /note="assembly-fragment"
 3541..4880
 /note="assembly-fragment"
 4981..6768
 /note="assembly-fragment"
 6869..8318
 /note="assembly-fragment"
 8419..10046
 /note="assembly-fragment"
 10147..12455
 /note="assembly-fragment"
 12556..15164
 /note="assembly-fragment"
 15265..16648
 /note="assembly-fragment"
 16749..29429
 /note="assembly-fragment"
 29530..41226
 /note="assembly-fragment"
 41327..57370

misc_feature /note="assembly_fragment"
57471..102968
misc_feature /note="assembly_fragment"
103069..164809
BASE COUNT 50889 a 32076 c 31689 g 48754 t 1401 others
ORIGIN

Query Match 20.0%; Score 340.8; DB 2; Length 164809;
Best Local Similarity 76.6%; Pred. No. 6e-57;
Matches 444; Conservative 0; Mismatches 132; Indels 4; Gaps 2;

QY 11 gaagctcttcacactacgctgtatgtcaactggtgagtcgccggcccatggtatgagaatt 70
DB 163517 GAGGAATTCCTCCGTGACCCCTGTCAGATGAGAGAGAGGCGCCCATTAAGAAAGAGT 163576
QY 71 gatcgaggatca--ataaagcttaattgaataataaagaaatatttcctt 128
DB 163577 GGGGTACAGATCAACACACACACTTAAGTGTATTAAGAAAGAAATTTTCT--CTT 163634
QY 129 tgaacttctccgtaagcaactgtgcctcctcttggtgggtcagtgltcacatcaat 188
DB 163635 TGAACCTTATCTGAGATCAGTCTTTTGTCTCCTGTGGGATCATTGTCACGCTCAAT 163694
QY 189 ggccttgaaggagctctgaagtcaggcttgagagcttgagattcagatgcttcaatc 248
DB 163695 GGCCTTACAGACCTCCTAGATCAAGTTGAGAGCCCTGGGAGATTCAGATCCTTCAGAT 163754
QY 249 ggttttcttccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 308
DB 163755 GGTTCCTTATTAATGTCAACGTCACTATGATATACATCAACTGAGAGACTTCGC 163814
QY 309 tgcagccattcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 368
DB 163815 AGCATTCATTAATGATCTCCGCTGCTGGGTCTATTAATGACAAATGACATATCCCTGA 163874
QY 369 taatgaacttgaatcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 428
DB 163875 CAATGACCCCTGGAGCCCTCAGCAGAGATGCCCTCCTGAGAAATCTCCATTCGACTC 163934
QY 429 aaatcgaagcagaagaagtgtcgtcgttctgtccatcccaagtgagccttctcact 488
DB 163935 AAATCTAGGCGCAGAGAGAGTGTCTGCTTGTCCATCCCCAGTGGAAAGCTCATTCATCT 163994
QY 489 gaatggagactatccagcagcaagtcgaagtcagacacgaacccctgtgtgagtgctgg 548
DB 163995 GAATGGAGCCTTCCCAACAGAGATGAGCCAGATACAGAGCCCTGTGTGATGGTGGGT 164054
QY 549 atatgatcaagactcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 588
DB 164055 ATATGACCAAGCTCTCCCTCCACCATGTGTGACTAAG 164094

Search completed: July 1, 2002, 15:55:28
Job time: 11742 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2002, 12:39:46 ; Search time 3305.87 Seconds
(without alignments)
5248.165 Million cell updates/sec

Title: US-09-674-235-10
Perfect score: 804
Sequence: 1 atggccttgaggagctctt.....acacaaagtgcacaaattt 804

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank1:
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_to:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_hg_hum:*
31: em_hg_inv:*
32: em_hg_other:*
33: em_hg_inv:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query Match Length DB ID Description

1	508	63.2	2684	6	AX074150	AX074150 Sequence
2	402	50.0	160945	2	AP003420	AP003420 Homo sapi
3	402	50.0	166848	2	AP002367	AP002367 Homo sapi
4	402	50.0	169776	9	AP001880	AP001880 Homo sapi
5	402	50.0	170682	2	AC073256	AC073256 Homo sapi
6	400.4	49.8	116273	2	AP000596	AP000596 Homo sapi
7	336.2	41.8	1566	9	BC022379	BC022379 Homo sapi
8	335.2	41.7	1846	9	AB062418	AB062418 Homo sapi
9	335.2	41.7	1986	6	AX179749	AX179749 Sequence
10	314.2	39.1	1977	6	AX074149	AX074149 Sequence
11	300	37.3	2597	10	RN0181	RN0181 Rattus norv
12	293.6	36.5	1995	10	AB051836	AB051836 Rattus norv
13	293.6	36.5	2253	9	AK057654	AK057654 Homo sapi
14	290.4	36.1	1964	10	BC021384	BC021384 Mus muscu
15	288.8	35.9	1911	10	RN013285	RN013285 Rattus no
16	277.6	34.5	2016	10	BC014805	BC014805 Mus muscu
17	277.6	34.5	2878	10	MMU132857	MMU132857 Mus muscu
18	275.6	34.3	116273	2	AP000596	AP000596 Homo sapi
19	275.6	34.3	160945	2	AP003420	AP003420 Homo sapi
20	275.6	34.3	164809	2	AC090778	AC090778 Homo sapi
21	269.6	33.5	169776	9	AP001880	AP001880 Homo sapi
22	269.6	33.5	2079	10	RN013285	RN013285 Rattus no
23	261.2	32.5	166848	2	AP002367	AP002367 Homo sapi
24	250	31.1	179934	9	AP000810	AP000810 Homo sapi
25	250	31.1	203598	2	AC090250	AC090250 Homo sapi
26	250	31.1	210723	2	AP001858	AP001858 Homo sapi
27	241.6	30.0	181234	2	AC103148	AC103148 Rattus no
28	241.6	30.0	185404	2	AC025409	AC025409 Homo sapi
29	238.8	29.7	207064	2	AC106680	AC106680 Rattus no
30	211.8	26.0	170682	2	AC073256	AC073256 Homo sapi
31	208.8	26.0	1737	6	AX191489	AX191489 Sequence
32	208.8	26.0	2455	6	AX230568	AX230568 Sequence
33	208.8	26.0	2665	6	AX191499	AX191499 Sequence
34	191.2	23.8	1779	10	AB005451	AB005451 Mus muscu
35	181	22.5	2516	9	AK055737	AK055737 Homo sapi
36	180.8	22.5	1653	6	AX191610	AX191610 Sequence
37	180.8	22.5	2047	6	AX230572	AX230572 Sequence
38	180.8	22.5	2194	6	AX191620	AX191620 Sequence
39	180.8	22.5	2210	6	E54995	E54995 Placental o
40	180.8	22.5	2210	9	AB026116	AB026116 Homo sapi
41	179.6	21.7	79141	2	AB050269	AB050269 Homo sapi
42	174.8	20.0	119151	2	AC099123	AC099123 Rattus no
43	160.4	20.0	120628	9	AP000655	AP000655 Homo sapi
44	160.4	20.0	153169	2	AC012153	AC012153 Homo sapi
45	160.4	20.0	153169	2	AP001092	AP001092 Homo sapi

ALIGNMENTS

RESULT 1
AX074150 AX074150 2684 bp DNA linear PAT 06-FEB-2001
LOCUS Sequence 6 from Patent WO0104283.
DEFINITION AX074150
ACCESSION AX074150
VERSION AX074150.1 GI:12710362
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2684)
AUTHORS Organic anion transporter genes and proteins
TITLE Patent: WO 0104283-A 6 18-JAN-2001,
JOURNAL Metabasis Therapeutics Inc. (US)
FEATURES
Location/Qualifiers
1..2684
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 800 a 533 c 541 g 810 t
ORIGIN

Query Match 63.2%; Score 508; DB 6; Length 2684;
 Best Local Similarity 100.0%; Pred. No. 2.4e-136;
 Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atggcctttgagagagctcttgatcaagttgagagccttgagagattcagatgctcat 60
 |||||
 Db 200 ATGGCTTTGAGAGAGCTCTTGATCAAGTTGGAGGCTTGGAGATTTCAGATGCTTCAT 259
 OY 61 ctggttttattcttccttccttcacatgltatataccctatctacgtcgaagaattc 120
 |||||
 Db 260 CTGGTTTATTCTTCTCCCTCTCATGTTATTATCCCTATATACCTGATGAGAACTTT 319
 OY 121 gctgcagccattccttgatcgttgctggttcacatgctgagacaataactgagct 180
 |||||
 Db 320 GCTGCAGCCATTCTCGTCATCGTCTGGTCCACATGCTGGACAAATATATGATCT 379
 OY 181 ggtatgaactggaactcctcagtgagaatgcccctcttgagaatcctatccactagac 240
 |||||
 Db 380 GGTATGAAAGCTGGAATCCTCAGTGAAGATGCCCTCTTGAATCTCTATCCCACTAGAC 439
 OY 241 tcaaatctgaagccagagaagtgctgcttgcttcacatcccaatgagactcttcac 300
 |||||
 Db 440 TCAATCTGAGGCCAGAGAGAGTGTCTCCCTTTGTCTATCCAGTGGAGCTTCTTAC 499
 OY 301 ctgaatggagactatcacagcacaagttagagcagacagaccctgltgagatgctg 360
 |||||
 Db 500 CTGAATGGGACTATCCACAGCACAAGTGAAGCAGACAGAACCCCTGTGGATGGCTGG 559
 OY 361 gatatgaactgaactccttccttcagacattgagcaaatggagacccctgtagatg 420
 |||||
 Db 560 GATATGATCAAGACTACTCTCCCTTGACCATGTGATCAAGTGGGACCTGTGATGTGAT 619
 OY 421 taccagtcactgaactcagtggttcaatcctactcctactcagtggaatgctggtgagagc 480
 |||||
 Db 620 TATCAGTCACTGAATCAAGTGGTCAATCTCTACCTTCTGACCTGATGATGCTGGAGAGC 679
 OY 481 atcatagtgagcattgcttcagacagct 508
 |||||
 Db 680 ATCATAGGTGGCCATGTCTCAGACAGGT 707

RESULT 2
 AP003420 160945 bp DNA linear PRI 28-AUG-2001
 LOCUS Homo sapiens genomic DNA, chromosome 11q, clone:RP11-614K12,
 DEFINITION complete sequence.
 ACCESSION AP003420
 VERSION AP003420.2 GI:15320509
 KEYWORDS HTG.
 SOURCE Homo sapiens DNA, clone:RP11-614K12.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 160945)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE Homo sapiens genomic DNA
 JOURNAL Published Only in Database (2001) In press
 REFERENCE 2 (bases 1 to 160945)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE Direct Submission
 JOURNAL Submitted (16-MAR-2001) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suenho-cho, Tsukuba, Ibaraki, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
 Tel:81-45-503-9111, Fax:81-45-503-9170)
 COMMENT On Aug 27, 2001 this sequence version replaced gi:13383333.
 FEATURES
 location/Qualifiers
 1..160945
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

/chromosome="11"
 /map="11q"
 /clone="RP11-614K12"
 BASE COUNT 46743 a 31738 c 32047 g 50417 t
 ORIGIN

Query Match 50.0%; Score 402; DB 9; Length 160945;
 Best Local Similarity 100.0%; Pred. No. 1.5e-105;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atggcctttgagagagctcttgatcaagttgagagccttgagagattcagatgctcat 60
 |||||
 Db 143970 ATGGCTTTGAGAGAGCTCTTGATCAAGTTGGAGGCTTGGAGATTTCAGATGCTTCAT 144029
 OY 61 ctggttttattcttccttccttcacatgltatataccctatctacgtcgaagaattc 120
 |||||
 Db 144030 CTGGTTTATTCTTCTCCCTCTCATGTTATTATCCCTATATACCTGATGAGAACTTT 144089
 OY 121 gctgcagccattccttgatcgttgctggttcacatgctgagacaataactgagct 180
 |||||
 Db 144090 GCTGCAGCCATTCTCGTCATCGTCTGGTCCACATGCTGGACAAATATATGATCTG 144149
 OY 181 ggtatgaactggaactcctcagtgagaatgcccctcttgagaatcctatccactagac 240
 |||||
 Db 144150 GGTATGAAAGCTGGAATCCTCAGTGAAGATGCCCTTTGAGATCTCTATCCCACTAGAC 144209
 OY 241 tcaaatctgaagccagagaagtgctgcttgcttcacatcccaatgagactcttcac 300
 |||||
 Db 144210 TCAATCTGAGGCCAGAGAGAGTGTCTCCCTTTGTCTATCCAGTGGAGCTTCTTAC 144269
 OY 301 ctgaatggagactatcacagcacaagttagagcagacagaccctgltgagatgctg 360
 |||||
 Db 144270 CTGAATGGGACTATCCACAGCACAAGTGAAGCAGACAGAACCCCTGTGGATGGCTGG 144329
 OY 361 gatatgaactgaactccttccttcagacattgagcaaatggagacccctgtagatg 402
 |||||
 Db 144330 GATATGATCAAGACTACTCTCCCTTGACCATGTGATCAAG 144371

RESULT 3
 AP002367 166848 bp DNA linear HTG 31-MAY-2000
 LOCUS Homo sapiens chromosome 11 clone RP11-151E18 map 11q, WORKING DRAFT
 DEFINITION SEQUENCE, 9 unordered pieces.
 ACCESSION AP002367
 VERSION AP002367.1 GI:8131631
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens DNA, clone:RP11-151E18.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 166848)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE Homo sapiens 166,848 genomic DNA of 11q
 JOURNAL Published Only in Database (2000) In press
 REFERENCE 2 (bases 1 to 166848)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE Direct Submission
 JOURNAL Submitted (29-MAY-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
 Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
 Tel:81-42-778-9924, Fax:81-42-778-9924)

----- Genome Center
 Center: RIKEN Genomic Sciences Center(GSC)
 Center code: RIKEN
 Web site: http://hgp.gsc.riken.go.jp/
 Contact: hattori@gsc.riken.go.jp
 ----- Project Information


```

misc-feature /note="assembly_fragment clone_end:T7 vector_side:left"
162066 .165410
/note="assembly_fragment clone_end:SP6 vector_side:left"
misc-feature 165511 .166848
/note="assembly_fragment"
BASE COUNT 55669 a 32391 c 31923 g 46064 t 801 others
ORIGIN

Query Match 50.0%; Score 402; DB 2; Length 166848;
Best Local Similarity 100.0%; Pred. No. 1.5e-105;
Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 atggccttgaggagcctctgtagtcaagttggaagccttggaagattcgaatgcttcat 60
|||||
Db 161305 ATGGCCTTTGAGGAGCTCTTGAGTCAAGTTGAGGCGCTTTGGAGATTTCAGATGCTTCAT 161364
|||||

Oy 61 ctggttttattcttcctctccatgattatataatccctcatatactgctgagaacttt 120
|||||
Db 161365 CTGGTTTTATTCTTCCTCTCCATGTTATTATATCCCTCATATATCTGCTGAGAACTTT 161424
|||||

Oy 121 gcttgagccatctcctgylcatcgttctggtggttccaatgctgagacaataactgagatc 180
|||||
Db 161425 GCTGCAGGCATTCCTGGTGCATCGTTGCTGGGTCCACATGCTGCAGCAATATATCTGAGATCT 161484
|||||

Oy 181 ggaatgaacatcgagatcctcagtgagaatgcccccttggaactctcatcccatagac 240
|||||
Db 161485 GGTAATGAACATGGAATCTCTCAGTCAATATCCCTCTTGGAATATCTATCCCATAGAC 161544
|||||

Oy 241 tcaatctgagagccagagaagtgtcgtgcttttgcattcccaagtcgagagccttccac 300
|||||
Db 161545 TCAAACTGAGGCGCAAGAGATGTGCTGCTTTGTCATCCCAAGTGGCAGGCTTCTCAC 161604
|||||

Oy 301 ctgaatgagcatctgcagacagacaaagtgaaggagacacagaaacccgtgtgagtgctg 360
|||||
Db 161605 CTGAATGGACTATCCACAGCACAAGTAGGAGGACAGACAGAACCCCTGTGGATGCTGG 161664
|||||

Oy 361 gtatagatcaaaactactctcccttgaccattgagctaaag 402
|||||
Db 161665 GTATATGATCAAAAGCTACTTCCCTTGACACATTGTGACTAAG 161706
|||||

RESULT 4
AP001880 169776 bp DNA linear PRI 26-APR-2001
LOCUS Homo sapiens genomic DNA, chromosome 11q, clone:CTD-3110P2,
DEFINITION complete sequences.
ACCESSION AP001880
VERSION AP001880.4 GI:13810518
KEYWORDS HTG.
SOURCE Homo sapiens DNA, clone:CTD-3110P2.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (sites)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens genomic DNA
Published Only in Database (2000) In press
2 (bases 1 to 169776)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (24-APR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suenho-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Apr 26, 2001 this sequence version replaced gi:10716820.
Location/Qualifiers
1. 169776
/organism="Homo sapiens"
/db_xref="taxon:9606"

```


106601 108529 contig of 1929 bp in length
 108630 110841 contig of 2212 bp in length
 110942 112367 contig of 1426 bp in length
 112468 113665 contig of 1198 bp in length
 113766 114974 contig of 1209 bp in length
 115075 116273 contig of 1199 bp in length

Sequence updated (26-May-2000).

* NOTE: This is a 'working draft' sequence. It currently consists of 29 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 38138: contig of 38138 bp in length
 38139 38238: gap of 100 bp
 38239 44698: contig of 6460 bp in length
 44699 44798: gap of 100 bp
 44799 49514: contig of 4716 bp in length
 49515 49614: gap of 100 bp
 49615 55534: contig of 5920 bp in length
 55535 55634: gap of 100 bp
 55635 60136: contig of 4502 bp in length
 60137 60236: gap of 100 bp
 60237 65623: contig of 5387 bp in length
 65624 65723: gap of 100 bp
 65724 70102: contig of 4379 bp in length
 70103 70202: gap of 100 bp
 70203 71387: contig of 1185 bp in length
 71388 71487: gap of 100 bp
 71488 73839: contig of 2352 bp in length
 73840 73939: gap of 100 bp
 73940 77366: contig of 3427 bp in length
 77367 77466: gap of 100 bp
 77467 80065: contig of 2599 bp in length
 80066 80165: gap of 100 bp
 80166 83825: contig of 3660 bp in length
 83826 83925: gap of 100 bp
 83926 86777: contig of 2852 bp in length
 86778 86877: gap of 100 bp
 86878 89783: contig of 2906 bp in length
 89784 89883: gap of 100 bp
 89884 92751: contig of 2868 bp in length
 92752 92851: gap of 100 bp
 92852 95068: contig of 2217 bp in length
 95069 95168: gap of 100 bp
 95169 96739: contig of 1571 bp in length
 96740 96839: gap of 100 bp
 96840 97912: contig of 1073 bp in length
 97913 98012: gap of 100 bp
 98013 99744: contig of 1732 bp in length
 99745 99844: gap of 100 bp
 99845 101153: contig of 1309 bp in length
 101154 101253: gap of 100 bp
 101254 102868: contig of 1615 bp in length
 102869 102968: gap of 100 bp
 102969 104921: contig of 1953 bp in length
 104922 105021: gap of 100 bp
 105022 106500: contig of 1479 bp in length
 106501 106600: gap of 100 bp
 106601 108529: contig of 1929 bp in length
 108530 108629: gap of 100 bp
 108630 110841: contig of 2212 bp in length
 110842 110941: gap of 100 bp
 110942 112367: contig of 1426 bp in length
 112368 112467: gap of 100 bp
 112468 113665: contig of 1198 bp in length
 113666 113765: gap of 100 bp
 113766 114974: contig of 1209 bp in length
 114975 115074: gap of 100 bp
 115075 116273: contig of 1199 bp in length.

Location/Qualifiers

source
 1...116273
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q13"
 /clone="XPac-662A14"
 1...38138
 /note="assembly-fragment"
 misc_feature
 38239..44698
 /note="assembly-fragment"
 misc_feature
 44799..49514
 /note="assembly-fragment"
 misc_feature
 49615..55534
 /note="assembly-fragment"
 misc_feature
 55635..60136
 /note="assembly-fragment"
 misc_feature
 60237..65623
 /note="assembly-fragment"
 misc_feature
 65724..70102
 /note="assembly-fragment"
 misc_feature
 70203..71387
 /note="assembly-fragment"
 misc_feature
 71488..73839
 /note="assembly-fragment"
 misc_feature
 73940..77366
 /note="assembly-fragment"
 misc_feature
 77467..80065
 /note="assembly-fragment"
 misc_feature
 80166..83825
 /note="assembly-fragment"
 misc_feature
 83926..86777
 /note="assembly-fragment"
 misc_feature
 86878..89783
 /note="assembly-fragment"
 misc_feature
 89884..92751
 /note="assembly-fragment"
 misc_feature
 92852..95068
 /note="assembly-fragment"
 misc_feature
 95169..96739
 /note="assembly-fragment"
 misc_feature
 96840..97912
 /note="assembly-fragment"
 misc_feature
 98013..99744
 /note="assembly-fragment"
 misc_feature
 99845..101153
 /note="assembly-fragment"
 misc_feature
 101254..102868
 /note="assembly-fragment"
 misc_feature
 102969..104921
 /note="assembly-fragment"
 misc_feature
 105022..106500
 /note="assembly-fragment"
 misc_feature
 106601..108529
 /note="assembly-fragment"
 misc_feature

Query Match 49.8%; Score 400.4; DB 2; Length 116273;
 Best Local Similarity 99.8%; Pred. No. 4.3e-105;
 Matches 401; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atggcctttagggagcctttagtcaagttagagccttggagatattcaatgcttcatt 60
 |||||
 Db 111822 ATGGCCTTTAGGAGCCTTTAGTCAAGTTCAGGCGCTTGGAGATTTCAGATGCTTCAT 111881

QY 61 ctggtttattcttcctcctcctcatgttatataccctacatactctagagaactt 120
 |||||
 Db 111882 CTGgttttttattcttcctcctcctcatgttatataccctacatactctagagaactt 111941

QY 121 gctgcagccattcctgctgcatcgttgcctgggtccacatgctggaacaataactgact 180
 |||||
 Db 111942 GCTgcagccattcctgctgcatcgttgcctgggtccacatgctggaacaataactgact 112001

QY 181 ggtatgaactgaatcctcagtgagaatgcctcttggagaattctataccactagac 240
 |||||
 Db 112002 GGTatgaactgaatcctcagtgagaatgcctcttggagaattctataccactagac 112061

QY 241 tcaaatctgagccagagaagtgtcgcttctgtccatcccaagtggcagcttctcac 300
 |||||||
 Db 112062 TCAAAATCTGAGCCAGAGAAAGTGTGCTTGTGTCATCCAGAGAGCTTCTTAC 112121
 |||||||
 QY 301 ctgaatggagactatcacagcagcaagtgtgagcagacagaaacccgtgtgtgagctgtg 360
 |||||||
 Db 112122 CTGAATGGAGACTATCCAGCAAGTGTGAGGACAGACAGAAACCCGTGTGTGATGGCTGG 112181
 |||||||
 QY 361 gtatatgacaaagctactcccttcgacattgtgactaaag 402
 |||||||
 Db 112182 GTATATGATCAAGACTACTCTCCCTTGACCATTTGTGACTAAG 112223
 |||||||
 RESULT 7
 BC022379 1566 bp mRNA linear PRI 04-FEB-2002
 LOCUS Homo sapiens, similar to ustr3, clone MGC:23972 IMAGE:4714598, mRNA,
 DEFINITION complete cds.
 ACCESSION BC022379
 VERSION BC022379.1 GI:18490377
 KEYWORDS MGC.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1566)
 AUTHORS Strausberg, R.
 JOURNAL Direct Submission
 Submitted (01-FEB-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: CLONTECH
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human
 Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mdickpaxl@stanford.edu
 Dickson, M., Schmitz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL Plate: 37 Row: d Column: 21
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.
 FEATURES
 source Location/Qualifiers
 1..1566
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:23972 IMAGE:4714598"
 /issue_type="Liver"
 /clone_lib="NIH_MGC_76"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"
 269..985
 /codon_start=1
 /product="Similar to ustr3"
 /protein_id="AAH2379.1"
 /db_xref="GI:18490378"
 /translation="MAFDLIGHAGDLMRFQLOTVLSIFAVATYLHMLNENFAFI
 PGRWCWHLINDNTVSDNDGALSODALRLISPLDSNMRPEKRFVHPQMLHLN
 GTPNWSADMEPCVDGWYDRISFSTSTYTEMDLVDOSLSVAKFVPMAGMAMGG
 IIGHSIDSSRNGNTIIPGHGNTIGNVPMFYCTIYDGRGFCRSLAYPPAGVCVILL
 CDLSIDKLARVCSVAHIQO"
 BASE COUNT 431 a 375 c 356 g 404 t
 ORIGIN

Query Match 41.8%; Score 336.2; DB 9; Length 1566;
 Best Local Similarity 78.1%; Pred. No. 1,86-86;
 Matches 404; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
 QY 1 atgagctttgagagagctcttgaatgaattgagagctctggagagattcagatgcttcat 60
 |||||||
 Db 269 ATGGCCTTTCAGGACCTCCGCGGTGACGCTGTGGAGACTGTGGAGATTCAGATTCCTTACG 328
 |||||||
 QY 61 ctggttttatcttcctccctcccaatgtaatacctataactcagagaaatt 120
 |||||||
 Db 329 ACTGTTTCTTCATCTTCTGCTGTTGCTACATFACCTTATTTATGCGGAGAACTTC 388
 |||||||
 QY 121 gctgcagccattcctggtatcgtgtggtgtgtccacatgctgagacaaataactgactc 180
 |||||||
 Db 389 ACTGATTCATFACCTGGCCATCGCTGCTGTGTCACATTCGAGACATGACACTGTCTCT 448
 |||||||
 QY 181 gctaatgaaactggaatccctcagtgaaagtccctctgtggaatctctatcccaactagc 240
 |||||||
 Db 449 GACAAATGACACTGGGGCCCTCAGCCAGATGACACTCTTGAAGAACTTCATCCCACTGGAC 508
 |||||||
 QY 241 tcaaatctgagccagagaagtgtcgcttctgtccatcccaagtggcagcttctcac 300
 |||||||
 Db 509 TCAAAATCTGAGCCAGAGAAAGTGTGCTTGTGTCATCCAGAGAGCTTCTTAC 568
 |||||||
 QY 301 ctgaatggagactatcacagcagcaagtgtgagcagacagaaacccgtgtgtgagctgtg 360
 |||||||
 Db 569 CTGAATGGAGACTATCCAGCAAGTGTGAGGACAGACAGAAACCCGTGTGTGATGGCTGG 628
 |||||||
 QY 361 gtatatgacaaagctactcccttcgacattgtgactaagtgtgagcctgtgtatgtgat 420
 |||||||
 Db 629 GTGATGACAGAAATCTCTTCTCTCATCCACATGCTGTGATGAGTGTGATGTGAC 688
 |||||||
 QY 421 tctagctcactgaaatcagatgtgtcattcttacttctgactgtgactgtgtgtgagagc 480
 |||||||
 Db 689 TCTCAATGACACTGACTTCTGAGGCTAATTTGTATTATTCATGCTGTGAAGATGATGGAGGC 748
 |||||||
 QY 481 atcataggtggccatgtctcagagagagtgctgtgtg 517
 |||||||
 Db 749 ATCTTAGCGGTCTATTATCAGACAGTAGCCGAGGTG 785
 |||||||
 RESULT 8
 AB062418 1846 bp mRNA linear PRI 05-JAN-2002
 LOCUS Homo sapiens ustr3 mRNA for huST3, complete cds.
 DEFINITION
 ACCESSION AB062418
 VERSION AB062418.1 GI:18148872
 KEYWORDS
 SOURCE Homo sapiens tissue_lib:fetal liver cdna to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Takanaga, H., Ohtsuki, S., Hosoya, K., and Terasaki, T.
 TITLE Isolation of novel clone of amphiphilic solute facilitator family
 from human fetal liver
 JOURNAL Unpublished
 2 (bases 1 to 1846)
 Takanaga, H., Ohtsuki, S. and Terasaki, T.
 Direct Submission
 Submitted (26-MAY-2001) Hitomi Takanaga, Tohoku University,
 Graduate School of Pharmaceutical Sciences, Aramaki, Aoba, Sendai,
 Aoba-ku, Miyagi 980-8578, Japan
 E-mail: takanaga@mai1.pharm.tohoku.ac.jp, Tel: 81-22-217-6832,
 Fax: 81-22-217-6886)
 FEATURES
 source Location/Qualifiers
 1..1846
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_lib="fetal liver"
 102..1763
 gene

CDS

/gene="ust3"
102..1763
/gene="ust3"
/codon_start=1
/product="HUST3"
/protein_id="BAB83517.1"
/db_xref="GI:18148873"
/translation="MAFDLGLHAGDLMPFOLIQVFLSIFAVATYLFHLENFTAFI
FGHRCVHILNDIVSDNDTGSALDRLISIPIDSNMRPEKCRREHPHOMLIHN
GTPMTSDADMEPCVDGWYDRISPSSTVTEMDLYCSOSLTSVAKFPMAGMYGG
ILGLHSDRGRFRFLRNCYLOVAIVGTALAPFLITCSFLSLGIDAMSLLITIT
MLLEWATHRRFQAGITIGMPCSGIAFMFLAGLAEIRWHLIQLVSVAPVFLTSL
SMLESARWLLINNKPERGLKELRKAHRSKGNKARDLTLLEIKSTMKKELEAAOK
KPSICEMLMHPNICKRISLSTFRANFMAVYGLNHYOHLGNVFLIQTLFGAVILL
ANCVAPMAKLYNNRASOMLMLFLAICLALIFVQEQMOTREVLAITGLGASALAN
TLAFHNGNEVITILIRAMGINATFANAGALAPLMITLSVSPPLMIYGVPPFI
SGFAFLIPGRNKPFLPDTIDOEKMRKDRPREKQDRPREVETQF"

BASE COUNT 458 a 472 c 420 g 496 t

ORIGIN

Query Match 41.7%; Score 335.2; DB 9; Length 1846;
Best Local Similarity 78.7%; Pred. No. 3.4e-86;
Matches 400; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 1 atggccttgaggagccttgatgaagttgagagccttgaggatctcagatgctcat 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102 ATGGCTTTCAGAGACCTCGTGCGTACGCTGGTGCAGATTCAGATTCCTTCAG 161
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 ctggttttattctccctctcctcatgttataatccctcattactgttagaactt 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 162 ACTGTTTTCTCTCAATCTTGCTGCTGTACATACCTTCATTTTATGCGAGAACTTC 221
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 gctgagcattcctcgtcatcgttgcgtgacacatgctgagcaataatactgact 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 222 ACTGATTCATATCCCTGGCCATCGCTGCTGCTGCATCTGAGCAATGACACTGTCTCT 281
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 ggtaatgaactggaactcctcagtagaagatgaccttgagaactctatccactagac 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 282 GACATATGACACTGGGCGCTCAGCAAGATGACATCTTGAGATCTCCATCCCACTGGAC 341
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 tcaaatctgaggagcagaagaagtgtgtgcttgccttgcacacccagtggaagctcttcac 300
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 342 TCACATATGAGGCGAGAAAGTGTGCTGCTTGTTCATCTCAGGAGAGCTCCCTTCAC 401
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 ctgaatggactatccacagacagaagtgaagcagaacccctgtgtgtagtgcg 360
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 402 CTGATATGGACCTTCCCAACAGTAGACGACATGAGAGCCCTGTGTGATGGCTGG 461
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 gtataatgacaagctacttccctcgaacattgtgactaagtggagcctgtgtatgat 420
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 462 GTGTATGACAGAAATCTCCTTCATCCACATCGTAGCTGAGATCGGATCTGGATGTGAC 521
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 tatcagtcacatgaatcagtggttcaattcctactctgacgcggaatgctgtgtgaggc 480
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 522 TCTCAATACACTGACTCTCAGGCTAAATTTGTATTATGCTGGAATGATGAGGAGGC 581
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 atcataagtgagcattgtctcagacagat 508
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 582 ATCCTAGCGGCTCATTTATTCAGACAGGT 609
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
AX179749 1986 bp DNA Linear PAT 06-AUG-2001
LOCUS AX179749
DEFINITION Sequence 37 from Patent W00146258.
ACCESSION AX179749
VERSION AX179749.1 GI:15132113
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1986)
AUTHORS Baughm, M.R., Butford, N., Au-Young, J., Lu, D.A., Yang, J., Reddy, R.,
Lal, P., Hillman, J.L., Azimzal, Y., Yue, H., Nguyen, D.B., Yao, M.G.,
Gandhi, A.R., Yang, Y.T., and Khan, F.A.
TITLE Transporters and ion channels
JOURNAL Patent: WO 0146258-A 37 28-JUN-2001;
Incyte Genomics, Inc. (US)
FEATURES
source location/Qualifiers
1..1986
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 3358383CB1"

BASE COUNT 505 a 496 c 457 g 528 t

ORIGIN

Query Match 41.7%; Score 335.2; DB 6; Length 1986;
Best Local Similarity 78.7%; Pred. No. 3.4e-86;
Matches 400; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 1 atggccttgaggagccttgatgaagttgagagccttgaggatctcagatgctcat 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 248 ATGGCTTTCAGAGACCTCGTGCGTACGCTGGTGCAGATTCAGATTCCTTCAG 307
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 ctggttttattctccctctcctcatgttataatccctcattactgttagaactt 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 308 ACTGTTTTCTCTCAATCTTGCTGCTGTACATACCTTCATTTATGCTGAGAACTTC 367
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 gctgagcattcctcgtcatcgttgcgtggtgacacatgctgagcaataatactgact 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 368 ACTGATTCATATCCCTGGCCATCGCTGCTGCTGCATCTGAGCAATGACACTGTCTCT 427
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 ggtaatgaactggaactcctcagtagaagatgaccttgagaactctatccactagac 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 428 GACATATGACACTGGGCGCTCAGCAAGATGACATCTTGAGATCTCCATCCCACTGGAC 487
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 tcaaatctgaggagcagaagaagtgtgtgcttgccttgcacacccagtggaagctcttcac 300
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 488 TCACATATGAGGCGAGAAAGTGTGCTGCTTGTTCATCTCAGTATGAGAGCTCCCTTCAC 547
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 ctgaatggactatccacagacagaagtgaagcagaacccctgtgtgtagtgcg 360
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 548 CTGATATGGACCTTCCCAACAGTAGACGACATGAGAGCCCTGTGTGATGGCTGG 607
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 gtataatgacaagctacttccctcgaacattgtgactaagtggagcctgtgtatgat 420
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 608 GTGTATGACAGAAATCTCCTTCATCCACATCGTAGCTGAGATCTGGATGTGAC 667
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 tatcagtcacatgaatcagtggttcaattcctactctgacgcggaatgctgtgtgaggc 480
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 668 TCTCAATACACTGACTCTCAGGCTAAATTTGTATTATGCTGGAATGATGAGGAGGC 727
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 atcataagtgagcattgtctcagacagat 508
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 728 ATCCTAGCGGCTCATTTATTCAGACAGGT 755
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
AX074149 1977 bp DNA Linear PAT 06-FEB-2001
LOCUS AX074149
DEFINITION Sequence 5 from Patent W00104283.
ACCESSION AX074149
VERSION AX074149.1 GI:12710361
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1977)
AUTHORS Sun, W.
TITLE Organic anion transporter genes and proteins
JOURNAL Patent: WO 0104283-A 5 18-JAN-2001;
Metabasis Therapeutics Inc. (US)

FEATURES	Location/Qualifiers
source	1..1977 /organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	502 a 491 c 450 g 534 t
ORIGIN	
Query Match	39.1%; Score 314.2; DB 6; Length 1977;
Best Local Similarity	77.3%; Pred. No. 4,3e-80;
Matches 395; Conservative	0; Mismatches 113; Indels 3; Gaps 1
OY	1 atgagcccttgaaagagctcttgatgcacgttggaagccttggaagattcagaatgctcat 60
Db	236 atggcctttcagagacctctctgggtacacgttggaaccttgagatgtccagatccttcag 295
OY	61 ctagttttatctctccctctcctatgattataatccctcatactgctagagaacttt 120
Db	236 actgtttttctctcaatcttttgcctgtctgtacatraccttcattttatgctggagaaacttc 355
OY	121 gctgaagccattccctgctgcatcgttctgctggtgcacacatgctgacataatactgattct 180
Db	356 actgcatttcattacctggccatgctgctgggtccacatncttgacacatgcactgctctt 415
OY	181 ggtatgaacatcggagaaacctcagatgaagatgacctctcttgagaatctctatcccaatagc 240
Db	416 gacatctgacactcggggcccttcagccaaagatgacactttgagaaatctccatccactggac 475
OY	241 tcaaatctgaggccgagagaagatgctgctgctcttgcatccccaatggagctcttcac 300
Db	476 tcaaacatgagagccagagaaagtgtgctgcttggttcattcctcactgacagctccttcac 535
OY	301 ctgaatgggactatccacagcaagaatgagcagacacagaacccctgtgtgatatgcttg 360
Db	536 ctgaatgggagaccttccccaacaaagtacagcagacatgagacccctgctggtgatggcttg 595
OY	361 gtaatagtcaagaactcttcccttcgacacat---tgtgactaaatggagactgattatgt 417
Db	596 gtgatatgacagaaatctcttctcattccaccatcggtgacctgaaatggagattggtatgt 655
OY	418 gattatcagtcactgaaatcagatggttcaactctcactcttgactggaatgctgtgtggga 477
Db	656 gacctctcaatcactgactcttcaatgagctgaaattgtgattcattcagatggctgtggga 715
OY	478 ggcatacatagtgccatgctatcacaagcgt 508
Db	716 ggcatctttaggcgcttattatcagacacaggt 746
RESULT 11	
RNDPTR LOCUS	RNDPTR 2597 bp mRNA linear ROD 30-MAR-1998
DEFINITION	Rattus norvegicus mRNA for putative integral membrane transport protein (UST1r).
ACCESSION	Y09945.1 GI:3004481
VERSION	Y09945
KEYWORDS	Integral membrane transport protein; UST1r gene.
SOURCE	Norway rat.
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1 (bases 1 to 2597) Schomayr, E., Splitzberger, F., Engelhardt, M., Martel, F., Oeding, N. and Grundemann, D.
AUTHORS	Molecular cloning and characterization of two novel transport proteins from rat kidney
TITLE	FEBS Lett. 425 (1), 79-86 (1998)
JOURNAL	66200080
MEDLINE	2 (bases 1 to 2597)
REFERENCE	Gruendemann, D.
AUTHORS	Direct Submission
TITLE	Submitted (10-DEC-1996) D. Gruendemann, University of Heidelberg,

FEATURES		Location/Qualifiers	
Source		1..2597	
gene		/organism="Rattus norvegicus"	
		/db_xref="taxon:10116"	
		/tissue_type="kidney"	
		87..1745	
		/gene="UST1r"	
		87..1745	
CDS		/gene="UST1r"	
		/codon_start=1	
		/product="putative integral membrane transport protein"	
		/protein_id="CAA71076.1"	
		/db_xref="GI:3004482"	
		/db_xref="SPTREMBL:O70609"	
		/translation="MAFDLNLQVGSLSGRFOILOMTFLINIIISPHLSLENFTAVY PNHRWVPIILDNDIVSGWNGNLSQDILLRIEIPDSLRPREKRRFVQPOMDLILNL GFSSVSEPDTEPCVDGNDVSDSTFLSTIIEMDVLVSSQSLDAKLEFLGLVNG ILYGLDPRFRRLILGLICASLOMATVEECAFAPFLICSLRFLAGISFTVENS LLIIMTRPRFOALATGLLCAAGATGVGLAFETVNMHHILHMSVPIFEFLVPP RMLSEARMLIMTNKLOKLEIKELIKVAIINKMNSDVLTEVARTIMELEASQTS SMDLEPRPNLRKRICLSLSEVRYVYLVGLINPRLHINWFLQCLEGITIPR NLVGFELVNHGRRISQFLFISLREISLIIAFPOEKQILRMVALTFGVSEVSAS SLVHNLNELPLIIRNALGVIAVIGLSTGAALSPLEMLRYSOPLWIIYGLVSLFIG GLVHLEPTEINQPIPADSIDQVEMGRASRQKQNDLIKVTQF"	
BASE COUNT	743 a 550 c 528 g 776 t		
ORIGIN			
Query Match	37.3%	Score 300;	DB 10; Length 2597;
Best Local Similarity	74.4%	Pred. No. 5.6e-76;	
Matches 378; Conservative	0;	Mismatches 130;	Indels 0; Gaps 0;
QY	1 atgagcccttgagagactctgagccaagtgtgagagccttgagagattcaatgcttat	60	
Db	87 ATGGCCTTCAGAGACTCCTGAAATCAAGTGGCAGCCTGGGAGGTTCCAGATCCTTAG	146	
QY	61 ctggttttattctctccctctcatgltatataatccctacatatactgtagaacttt	120	
Db	147 AAGACTTTTCATTTGATCTTCACATATAATATTTCTCCATTCATTCATGAGAACTTC	206	
QY	121 gctgagccattccctgfgcatcgtgtgctgggtgccacatgctggagacaataactgatact	180	
Db	207 ACAGCAGTCAATCCCAATCATCATGCTGGTGGGCCCATCTTCGACAAATGACACTGTCCT	266	
QY	181 ggtaaatgaactggaatcctcagtaagaatgccccttgagaatcctatcccaatagac	240	
Db	267 GGTATATGACAAATGGAGACCTGAGCCAAATGACCTCCGAGAGGTCTATATCCCTCGAC	326	
QY	241 tcaatctgagagccagagaagtgtcgtgcttgttgcataccccaagtggcagcttctaac	300	
Db	327 TCCGATCGAGACCGAGAAAGTGTGTGGCTTTGCCAACACAGAGTGGAGATCTCTTCAT	386	
QY	301 ctggaatggactatccagacaagaatgtagcagaacagacacctgtgtgatatggctgg	360	
Db	387 CTTAAACGGTACTCTTTCCAGTGTGACAGAGCCACAGACAGACCCCTGTGTGATGGCTGG	446	
QY	361 gatatgataaagctacttccctcgaccatltgtaactagtgagtgagactgtagat	420	
Db	447 GTGTATGACCAAGACACTTCTTCTTCCACATATTACCGAGTGGAGCTGTTGGAA	506	
QY	421 tataatgtacccaagaatcagtggttcaatctctacttctgactggaatgctgtgtggagac	480	
Db	507 TCTGAGTACAGGATCAATAGCTAAGTTTCTATTCTTGACTGCTATCTTAGTGGAAAT	566	
QY	481 atcataagtgagcatgtctcagacaggt	508	
Db	567 ATCCATATATGGCCCTTACACAGACAGGT	594	

Db 264 GATAATGACCTAGATGCTGAGCCAGATGACCTCTGAGAGTCTCCATCCCCCTGGAC 323
 QY 241 tcaaatctgaagccagagaagtgctgctgtgctccatcccaagtgagcttctcac 300
 Db 324 TCCAAACCTGAGACCGGATTAATGCGCTATATTCACACACAGTGCGATCTCTTCAT 383
 QY 301 ctgaatggactatcccaagcagcagtgagagcagcagcagcagcagcagcagcagc 360
 Db 384 CTGAATGGACCTTCCCTCCAGCTGTAACAGCCAGACACAGAGCCCTGTGTGGATGGCTGG 443
 QY 361 gataatgaatcaagctacttccctcagccatgtgactaagtgaggagctgtatgtat 420
 Db 444 GTGATGACCAAGACGACCTCTCTCCACACAGTACACCCAGTGAGGAGCTGTGTGGA 503
 QY 421 tacaatctgaagcagcagtgagtgctcaatctcactctgactggaatgctgtgtgaggc 480
 Db 504 TCTCAGGACCTGAATTCAGTGTGCTTAATTAATACATGATGCTGTATTAATGATAT 563
 QY 481 atcataagtgccatgctcagacagagt 508
 Db 564 ATCATGGGTGGCCACTTATCAGACAGT 591

RESULT 15
 RNO132859 1911 bp mRNA linear ROD 04-JAN-2002
 LOCUS Rattus norvegicus mRNA for putative integral membrane transport protein UST4r.
 DEFINITION
 ACCESSION AJ132859.1 GI:18076048
 VERSION AJ132859.1
 KEYWORDS integral membrane transport protein; UST4r gene.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (sites)
 REFERENCE 1 (sites)
 AUTHORS Honold,C., Gruendemann,D. and Schoemig,E.
 JOURNAL Unpublished
 TITLE 2 (bases 1 to 1911)
 AUTHORS Gruendemann,D.
 TITLE Direct Submission
 JOURNAL Submitted (08-FEB-1999) Gruendemann D., Dept. of Pharmacology, University of Heidelberg, INF 366, 69120 Heidelberg, 69120, GERMANY
 FEATURES
 source
 1..1911
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 /tissue_type="liver"
 121..1779
 /gene="UST4r"
 121..1779
 /gene="UST4r"
 /codon_start=1
 /product="putative integral membrane transport protein"
 /protein_id="CAC79639.1"
 /db_xref="GI:18076049"

gene
 CDS
 BASE COUNT 535 a 395 c 397 g 584 t
 ORIGIN
 Query Match 35.9%; Score 288.8; DB 10; Length 1911;
 Best Local Similarity 73.0%; Pred. No. 1e-72;

Matches 371; Conservative 0; Mismatches 137; Indels 0; Gaps 0;
 QY 1 atggccttgaggagcctcttgatcaagtggagagccttggaagttccatgcttcat 60
 Db 121 ATGGCCTTTGAGGAGACCTTAATCAGGTAGAGGCGCTAGGAGATCCAGATCCTTCAG 180
 QY 61 ctggttttattcttccctccatcagtgatataatccctatatagctgtagaactt 120
 Db 181 ATGCTTTTGTGCTTCCACAGTGTATGTGTAAGTCAATATATATAGAGAACTT 240
 QY 121 gctgagccatctcctgctacgtgtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 180
 Db 241 ACTGACCCATTTCCAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 QY 181 ggtatgaactggaatctcagtgagagtgccctctgagagatctcatccactagac 240
 Db 301 GATTAACGACGTAGATTTCTAACCAGATTAATCTCTGAAGATCTCCATCCCTTGGAC 360
 QY 241 tcaaatctgaagccagagaagtgctgctgtgcttgcacatcccaagtgagccttctcac 300
 Db 361 TCCACCTGAGACAGATTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 301 ctgaatggactatcccaagcagcagtgagagcagcagcagcagcagcagcagcagc 360
 Db 421 TTGAATGATTAATTTCTCCAGTATTAATTAATTAATTAATTAATTAATTAATTAAT 480
 QY 361 gataatgaatcaagctacttccctcagccatgtgactaagtgaggagctgtatgtat 420
 Db 481 GTATATGACAGAGACGACATTTTCATTCACACTGAGACTGAGACTGAGACTGAGACTG 540
 QY 421 tacaatctgaagcagcagtgagtgctcaatctcactctgactggaatgctgtgtgaggc 480
 Db 541 TCTCAGGACCTGAATTCGACACTTAATTTATGATGATGATGATGATGATGATGATGAT 600
 QY 481 atcataagtgccatgctcagacagagt 508
 Db 601 ATGTAAATGGCTTCTGTCTCAGACAGT 628

Search completed: July 1, 2002, 15:41:50
 Job time: 10924 sec

EMBL: 297028; CAB09724.1; -
 DR InterPro: IPR003662; sub-transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 FT Transmembrane; Transport; Glycoprotein; Ion transport.
 FT TRANSMEM 148 168 POTENTIAL.
 FT TRANSMEM 175 195 POTENTIAL.
 FT TRANSMEM 206 226 POTENTIAL.
 FT TRANSMEM 261 281 POTENTIAL.
 FT TRANSMEM 409 429 POTENTIAL.
 FT TRANSMEM 496 516 POTENTIAL.
 FT CAROHD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 562 AA; 61992 MW; 7DC1A67F32801D2D CRC64;

Query Match 31.2%; Score 439; DB 13; Length 562;
 Best Local Similarity 27.4%; Pred. No. 5.4e-34;
 Matches 99; Conservative 54; Mismatches 84; Indels 124; Gaps 5;

QY 1 MAFELLISQVGLGRFQMLHLVFLIPSLMLIPHLLENFAAIPGHRGVHMLDNTGS 60
 DB 1 MFESELDKQVSGRQVQLQVPMVAPRLMMASHNTLQNTAIPAHNC-----RPPAN 54
 QY 61 GNETGISLSDALLRISIPDLSNLRPEKCRFPVHPQWOLHLNGTISTEADT----- 113
 DB 55 LQQAFLSLSESL-ITVPLDAGKRPORRYAAPQWHLGKNGTSGSLDADTESMDAA 113
 QY 114 -EPYDGVNVDOSYSPSTIVTKMDLVCDYSLKSVQFLLTGMVGGIIGHVSDR-- 169
 DB 114 LQESDGSVSYSTVSTYSLISEMHLVCDMHSFKQGTIYMGVLLVGLLFGGLSDRYGR 173
 QY 170 ----- 169
 DB 174 RILLISNLAVSGTCAAFSSSSSLFCVFRFGGLAISGLNTFSLIVEMIPRIIRA 233
 QY 170 ----- 184
 DB 234 VGTTCYCTIGQLLVLVLAFFIRDMRWLTLAVSLPFVFFLIAMWFHSSRWLMSRT 293
 QY 185 DEGLKARKVARTNGIKNAEETLIEVVRSTMOEELDAQRTKYVCDLFRNDSMKR-IC 243
 DB 294 EHALNKLKSVARENGRHEAEKLIKMLHESMKKEMSCDGSYSILDFNTPAMKRRLC 353
 QY 244 I 244
 DB 354 I 354

RESULT 13
 ID 061185 PRELIMINARY; PRT; 545 AA.
 AC 061185;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE KIDNEY-SPECIFIC TRANSPORT PROTEIN.
 GN SLC22A6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX MEDLINE=97197822; PubMed=9045672;
 RA Lopez-Nieto C.E., You G., Bush K.T., Barros E.J., Beier D.R.,
 Nigam S.K.;
 RT "Molecular cloning and characterization of NKT, a gene product related
 to the organic cation transporter family that is almost exclusively
 expressed in the kidney";
 RT J. Biol. Chem. 272:6471-6478(1997).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

-1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC EMBL: U52842; AAC53112.1; -
 DR MGD: MGI:892001; SLC22a6
 DR InterPro: IPR003662; sub-transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Transmembrane.
 SQ SEQUENCE 545 AA; 59983 MW; CD3B762E5984933A CRC64;

Query Match 31.2%; Score 438.5; DB 11; Length 545;
 Best Local Similarity 28.7%; Pred. No. 5.8e-34;
 Matches 102; Conservative 52; Mismatches 77; Indels 125; Gaps 6;

QY 1 MAFELLISQVGLGRFQMLHLVFLIPSLMLIPHLLENFAAIPGHRGVHMLDNTGS 60
 DB 1 MAFNDLLKQVSGRQVQLQVPMVAPRLMMASHNTLQNTAIPAHNC-----RPPAN 54
 QY 61 GNETGISLSDALLRISIPDLSNLRPEKCRFPVHPQWOLHLNGTISTEADTPECVDM 120
 DB 55 AN-----LSKDGGLFARLPIDKGRPEKCLFPPH-----NCT-EANGTGTPECLDGM 103
 QY 121 VYDOSYEPSTIVTKMDLVCDYSLKSVQFLLTGMVGGIIGHVSDR----- 169
 DB 104 VIDNSTFSTIVTEWMLVCSHRAFRQLAQLFMVGVLLGAMFGYLDRLGRKKVLLNY 163
 QY 170 ----- 171
 DB 164 LQAVSGTCAAVAPNTVYCIIRLLSGMSLSTAINCMILNEMMIFHRAVYGLIGYV 223
 QY 172 ----- 192
 DB 224 YSLGQFLAGIYAVPMMHNLQAVSPPEVAFIYSWPFIESARWYSSSGRLDLTLRALQ 283
 QY 193 KYARTNGIKNAEETLIEVVRSTMOEELDAQRTKYVCDLFRNDSMKRKLCTIVFL 248
 DB 284 RVARINGQDEGAKLSIEVLQTSLOKELFLNKGASAMELHCPILURLCLISML 339

RESULT 14
 ID 096TC1 PRELIMINARY; PRT; 542 AA.
 AC 096TC1;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE ORGANIC ANION TRANSPORTER 3.
 GN HOAT3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cha S., Sekine T., Kanai Y., Endou H.;
 RT "Molecular cloning and characterization of human organic anion
 transporter 3";
 RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB042505; BAB47393.1; -
 SQ SEQUENCE 542 AA; 59857 MW; 5053B36439554CCC CRC64;

Query Match 27.8%; Score 390.5; DB 4; Length 542;
 Best Local Similarity 25.4%; Pred. No. 2.6e-29;
 Matches 91; Conservative 46; Mismatches 70; Indels 151; Gaps 5;

QY 1 MAFELLISQVGLGRFQMLHLVFLIPSLMLIPHLLENFAAIPGHRGVHMLDNTGS 60
 DB 1 MTFSELDKQVSGRQVQLQVPMVAPRLMMASHNTLQNTAIPAHNC-----RPPAN 54
 QY 51 VHMIDNNTGSGNETGISLSDALLRISIPDLSNLRPEKCRFPVHPQWOLHLNGTISTSE 110
 DB 61 V-----LPMGPNGKPERCLRFVHP-----MSLPTDQ 89

RT Expression cloning and characterization of a novel multispecific
 RT organic anion transporter.";
 RL J. Biol. Chem. 272:18526-18529(1997).
 CC -1- FUNCTION: MEDIATES THE NA(+)-INDEPENDENT TRANSPORT OF ORGANIC
 CC ANIONS SUCH AS CYCLIC NUCLEOTIDES, PROSTAGLANDIN E2 AND URIC ACID.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. BASOLATERAL
 CC MEMBRANE (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: STRONG EXPRESSION IN KIDNEY. VERY WEAK
 CC EXPRESSION IN BRAIN. NOT DETECTED IN HEART, LUNG, LIVER, SPLEEN
 CC SKELETAL MUSCLE, SMALL INTESTINE, LARGE INTESTINE, EYE OR TESTIS.
 CC EXPRESSED IN THE PROXIMAL TUBULE IN THE KIDNEY.
 CC -1- PTM: GLYCOSYLATED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SLC22 FAMILY OF TRANSPORTERS.
 DR EMBL; AF008221; AAC18772.1; -;
 DR EMBL; AB004559; BAA22086.1; -;
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Transmembrane; Transport; Glycoprotein; Ion transport.
 FT TRANSMEM 136 156 POTENTIAL.
 FT TRANSMEM 196 216 POTENTIAL.
 FT TRANSMEM 249 269 POTENTIAL.
 FT TRANSMEM 338 358 POTENTIAL.
 FT TRANSMEM 396 416 POTENTIAL.
 FT TRANSMEM 485 505 POTENTIAL.
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 551 AA; 60766 MW; 8BA47BE628324BF2 CRC64;

Query Match 32.5%; Score 457.5; DB 11; Length 551;
 Best Local Similarity 29.2%; Pred. No. 8.5e-36;
 Matches 104; Conservative 52; Mismatches 81; Indels 119; Gaps 5;

QY 1 MAFELLISQVGGIGRQMLHVLFIPLSLMLIPHLLLENFAAIPGRCVHMLDNNNGS 60
 DB 1 MAFNDLLKQVGGVGRQRIQVTVLVPLMLASHNTLQNTAIPPHHC-----RPPAN 54
 QY 61 GNETGILSEDALLRISIPDLSNLRPEKCRFRVHPQQLHNTGISTSEADTEPCVDGW 120
 DB 55 AN-----LSKDGLQVAVPFWHRYLQLVSPFPFAFVYSWPFIESARWYSTPGRDLTLTKALQ 109
 QY 121 VYDQSYFPSTIVTKMDLVCDYQSLKSVQFLLTGMLVGGIIGHSYDR-----169
 DB 110 VYDNSTFPSTIVTEWMLVCSHRAFRQLAQSIVMGVLLGAMVFGYIADRLGRKKVILANT 169
 QY 170 -----WL-----171
 DB 170 LGTAVSGTCAAPNVTYCVFRLLSGMSIASIANGMTLVNEMMPIHTRAYVGTLAGYV 229
 QY 172 -----VESARWILITNKUDEGLKALR 192
 DB 230 YSLGQPLAGIYAVPFWHRYLQLVSPFFIATFYSWPFIESARWYSSSGRGDLTLTKALQ 289
 QY 193 KVARINGIKNAEETLNEVYSTMOEELDAQTKTVCDLFRNPSRKRCICLVFL 248
 DB 290 RVARINGKEGAKLSIEVLRTSLQKELTSLKQASAMELLRCPTLRHLFLCLSL 345

RESULT 11
 Q9TSY7 PRELIMINARY; PRT; 551 AA.
 AC Q9TSY7;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE RENEL ORGANIC ANION TRANSPORTER 1 (RBOAT1).
 GN ROAT1.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY CORTEX;
 RC Bahn A., Knabe M., Hillmann A., Burckhardt G.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL; A2242871; CAB62587.1; -;
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Transmembrane.
 SO SEQUENCE 551 AA; 60588 MW; 69C25F96B154517E CRC64;

Query Match 32.0%; Score 449.5; DB 6; Length 551;
 Best Local Similarity 29.5%; Pred. No. 5.1e-35;
 Matches 105; Conservative 49; Mismatches 83; Indels 119; Gaps 4;

QY 1 MAFELLISQVGGIGRQMLHVLFIPLSLMLIPHLLLENFAAIPGRCVHMLDNNNGS 60
 DB 1 MAFNDLLKQVGGVGRQRIQVTVLVPLMLASHNTLQNTAIPPHHC-----RPPAN 56
 QY 61 GNETGILSEDALLRISIPDLSNLRPEKCRFRVHPQQLHNTGISTSEADTEPCVDGW 120
 DB 57 -----LSKDGLQVAVPFWHRYLQLVSPFPFAFVYSWPFIESARWYSTPGRDLTLTKALQ 109
 QY 121 VYDQSYFPSTIVTKMDLVCDYQSLKSVQFLLTGMLVGGIIGHSYDR-----169
 DB 110 VYDNSTFPSTIVTEWMLVCSHRAFRQLAQSIVMGVLLGAMVFGYIADRLGRKKVILANT 169
 QY 170 -----WL-----171
 DB 170 LGTAVSGTCAAPNVTYCVFRLLSGMSIASIANGMTLVNEMMPIHTRAYVGTLAGYV 229
 QY 172 -----VESARWILITNKUDEGLKALR 192
 DB 230 YSLGQPLAGIYAVPFWHRYLQLVSPFPFAFVYSWPFIESARWYSTPGRDLTLTKALQ 289
 QY 193 KVARINGIKNAEETLNEVYSTMOEELDAQTKTVCDLFRNPSRKRCICLVFL 248
 DB 290 KVARINGKEGAKLSIEVLRTSLQKELTSLKQASAMELLRCPTLRHLFLCLSL 345

RESULT 12
 O57379 PRELIMINARY; PRT; 562 AA.
 AC O57379;
 DT 01-JUN-1998 (Tremblrel. 06, Created)
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE RENEL ORGANIC ANION TRANSPORTER.
 OS Pseudopleuronecta americanus (Winter flounder).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Pleuronectiformes;
 OC Pleuronectidae; Pleuronectidae; Pseudopleuronectes.
 OX NCBI_TaxID=8265;
 RN [1]
 RP SEQUENCE FROM N.A., AND FUNCTION.
 RC TISSUE=KIDNEY;
 RC MEDLINE=98072393; PubMed=9409735;
 RA Wolf N.A., Werner A., Burkhardt S., Burckhardt G.;
 RT transporter cloning and characterization of a renal organic anion
 RT transporter from winter flounder.";
 RL FEBS Lett. 417:287-291(1997).
 CC -1- FUNCTION: MEDIATES THE NA(+)-INDEPENDENT TRANSPORT OF ORGANIC
 CC ANIONS SUCH AS P-AMINOHIPPURATE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. BASOLATERAL
 CC MEMBRANE (POTENTIAL).
 CC -1- PTM: GLYCOSYLATED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SLC21 FAMILY OF TRANSPORTERS.

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC TISSUE=KIDNEY.
 RX MEDLINE=99103997; PubMed=9887087;
 RA Hosoyamada M., Sekine T., Kanai Y., Endou H.;
 RT "Molecular cloning and functional expression of a multispecific
 organic anion transporter from human kidney.";
 RL Am. J. Physiol. 276:F122-F128(1999).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=98433806; PubMed=9762842;
 RA Reid G., Wolff N.A., Dautzenberg F.M., Burckhardt G.;
 RT "Cloning of a human renal P-aminohippurate transporter, hROAT1,";
 RL Kidney Blood Press. Res. 21:233-237(1998).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RX MEDLINE=99137667; PubMed=9950961;
 RA Lu R., Chan B.S., Schuster V.L.;
 RT "Cloning of the human kidney PAH transporter: narrow substrate
 specificity and regulation by protein kinase C.";
 RL Am. J. Physiol. 276:F295-F303(1999).
 RN [4]
 RN SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RX MEDLINE=99160894; PubMed=10049739;
 RA Race J.E., Grassl S.M., Williams W.J., Holtzman E.J.;
 RT "Molecular cloning and characterization of two novel human renal
 organic anion transporters (hOAT1 and hOAT3).";
 RL Biochem. Biophys. Res. Commun. 255:508-514(1999).
 RN [5]
 RN SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RX MEDLINE=99393620; PubMed=10462545;
 RA Chihlar T., Lin D.C., Pritchard J.B., Fuller M.D., Mendel D.B.,
 RA Sweet D.H.;
 RT "The antiviral nucleotide analogs cidofovir and adefovir are novel
 substrates for human and rat renal organic anion transporter 1.";
 RL Mol. Pharmacol. 56:570-580(1999).
 RN [6]
 RN SEQUENCE FROM N.A.
 RA Bahn A., Prawitt D., Reid G., Enklaar T., Wolff N.A., Hillemann A.,
 RA Godehardt S., Butler D., Knabe B., Schulten H.J., Gunawan B.,
 RA Fueszel L., Zabel B., Burckhardt G.;
 RT "Genomic cloning and characterization of the human renal organic anion
 transporter gene (hOAT1).";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MEDIATES THE NA(+)-INDEPENDENT TRANSPORT OF ORGANIC
 ANIONS SUCH AS P-AMINOHIPPURATE AND ALPHA-KETOGLUTARATE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. BASOLATERAL
 MEMBRANE (POTENTIAL).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: ISOFORM hOAT1-1 (SHOWN HERE)
 AND ISOFORM hOAT1-2; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN KIDNEY. ALSO DETECTED IN
 BRAIN AND AT LOW LEVELS, IN SKELETAL MUSCLE AND PLACENTA. NOT
 EXPRESSED IN HEART, LUNG, LIVER, PANCREAS, COLON, THYMUS, SMALL
 INTESTINE OR PERIPHERAL BLOOD LEUKOCYTES. IN THE KIDNEY, STRONGLY
 EXPRESSED IN PROXIMAL TUBULE CELLS.
 CC -1- PTM: GLYCOSYLATED.
 CC -1- SIMILARITY: BELONGS TO THE SLC22 FAMILY OF TRANSPORTERS.
 DR EMBL: AF057039; AAC70004.1; -;
 DR EMBL: AB009697; BAA75072.1; -;
 DR EMBL: AF104038; AAD10052.1; -;
 DR EMBL: AB009698; BAA75073.1; -;
 DR EMBL: AF097490; AAD19356.1; -;
 DR EMBL: AF124373; AAD55356.1; -;
 DR EMBL: AJ249369; CAB77184.1; -;
 DR InterPro: IPR003662; sub.transporter.
 DR Pfam: PF00083; sugar_tr.1.
 KM Transmembrane; Transport; Alternative splicing; Ion transport;
 Glycoprotein.

FT TRANSMEM 136 156 POTENTIAL.
 FT TRANSMEM 185 205 POTENTIAL.
 FT TRANSMEM 249 269 POTENTIAL.
 FT TRANSMEM 338 358 POTENTIAL.
 FT TRANSMEM 396 416 POTENTIAL.
 FT TRANSMEM 426 446 POTENTIAL.
 FT TRANSMEM 485 505 POTENTIAL.
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPPLIC 523 535 MISSING (IN ISOFORM OAT1.2).
 FT CONFLICT 14 14 G->S (IN REF. 3).
 FT CONFLICT 563 563 L->F (IN REF. 2).
 SQ SEQUENCE 563 AA; 61816 MW; 74AD3EA2678032E4 CRC64;
 Query Match 32.6%; Score 458.5; DB 4; Length 563;
 Best Local Similarity 29.2%; Pred. No. 7e-36;
 Matches 104; Conservative 51; Mismatches 82; Indels 119; Gaps 5;
 QY 1 MAPEELLQGVGGIGRROMLHVFLPILSLMLIPHIILENFAAIPGRWCWYHMDNNTGS 60
 DB 1 MAPNDLQGVGGVGRQOQVTLVPLMLASHNTLQNTAIPTHHC-----RRPAD 54
 QY 61 GNETGLISDALIRISIPLDNLNRPKCRFVHPQWOLHNGTINSTSEADTEPCVDGW 120
 DB .55 AN-----LSKNKGLVWLPRDQGPESCLRFSPQWMLPFLNGT-EANGTGATPCPDGW 109
 QY 121 VYDOSFPSTIYTKWPLVCDYQSLKSVYQFLITGMVGGIIGHSVDR----- 169
 DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 169
 QY 110 IYDNTFPTIYEMDVLGSHRALROSLVMVGLGAMVGYLRLGRKRLILNY 169
 QY 170 -----WL----- 171
 DB 170 LQTVASGTCAAFAPNPPIYCAFRLLSGMALAGISLNCMTLNVEMPHITACVGTILGY 229
 QY 172 -----VESARWLLITNKLDEGLRALR 192
 DB 230 YSLGQFLAGVAVPWHRLQLLVSAFFAFIYSWFIESARWSSSGRLDTLRALQ 289
 QY 193 KVARNGIKNAEETLIEVYRSTMOEELDAQTKTYCDLFRNSMKRRCILVFL 248
 DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 248
 DB 290 RVARINKREBAGKLSMEVLRASLOKELTMGKGQASAMELLRQPTLRHFLPLCSML 345
 RESULT 10
 ID 035956 PRELIMINARY; PRT; 551 AA.
 AC 035956;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
 DE RENAL ORGANIC ANION TRANSPORT PROTEIN 1.
 GN SLC22A6 OR ROAT1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RN SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
 RC TISSUE=KIDNEY;
 RX MEDLINE=98043701; PubMed=9374486;
 RA Sweet D.H., Wolff N.A., Pritchard J.B.;
 RT "Expression cloning and characterization of ROAT1. The basolateral
 organic anion transporter in rat kidney.";
 RL J. Biol. Chem. 272:30088-30095(1997).
 RN [2]
 RN SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
 RP STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
 RX MEDLINE=97373539; PubMed=9228014;
 RA Sekine T., Watanabe N., Hosoyamada M., Kanai Y., Endou H.;

```

RESULT 7
G9NOC2
ID 09NOC2 PRELIMINARY; PRT; 506 AA.
AC 09NOC2;
DT 01-OCT-2000 (TREMBLrel, 15, Created)
DT 01-OCT-2000 (TREMBLrel, 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel, 19, Last annotation update)
DE PURATIVE ORGANIC ANION TRANSPORTER.
GN OAT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RX TISSUE=KIDNEY;
MEDLINE=20422308; PubMed=10964714;
RA Bahh A., Prawitt D., Buttler D., Reid G., Enklaar T., Wolff N.A.,
RA Ebdlinghaus C., Hillemann A., Schulten H.J., Gunawan B., Fuenesi L.,
RA Zebel B., Burckhardt G.,
RT 'Genomic structure and in vivo expression of the human organic anion
RT transporter 1 (hOAT1) gene.';
RL Biochem. Biophys. Res. Commun., 275:623-630(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL: AI251529; CAB94830.1; -.
DR InterPro: IPR003662; sub:transporter.
DR Pfam: PF00083; sugar_tr; 1.
DR Transmembrane.
SQ SEQUENCE 506 AA; 55858 MW; D8EBAEBA113E6C5E CRC64;

```

Query Match	32.6%;	Score 458.5;	DB 4;	Length 506;
Best Local Similarity	29.2%;	Pred. No. 6.1e-36;		
Matches 104; Conservative	51;	Mismatches 82;	Indels 119;	Gaps
Oy	1	MAFEELLQVGGIGRGFOMLHVEFLIPSLMLLPIHLLENFNAAPADGRHCWMMLDNNTGS	60	
Dd	1	MAFNLLIQGVGVRGQQIQYTLVVLPLLMASHHTLOFTAAIPTTHNC-----RPPAD	54	
Oy	61	GNETIIEEDALLRIISTIPDSNLNRPEKCRRFVHPQOWMLHTNGTSHSSEADTEPCWDG	120	
Dd	55	AN----LSKNGLGLEWLPFRDROGPESCLREFTSPQWGLPENG--EANGCATEPTCTDGW	109	
Oy	121	VYDOSYFPSTIVTKMDVDCPDQOSKSVAQVQLLTGMLYGSGITGVSDR-----	168	
Dd	110	IYDNSTFPTIVTEMDLVCSHRALROLQAOSTLYMGVLLGAMVEGYLADRGRKKVILTY	169	
Oy	170	-----WL-----	171	
Dd	170	LQTAVSGTCAFAPNPETCYCAFRLSGMALAGISLNCMTLVNEMPHTRACYGTLLGIY	229	
Oy	172	-----VESARWLITTNKLDEGALKALR	192	
Dd	230	YSLGQFLLAGYAAYAVPWHRHLQLLVSAFFPAFFITYSWFFIESARHMSSSRGLDTLRALQ	289	

RESULT	8	
ID	09N0A6	PRELIMINARY; PRT; 519 AA.
AC	09N0A6;	
DT	01-OCT-2000 (TREMblrel. 15, Created)	
DT	01-OCT-2000 (TREMblrel. 15, last sequence update)	
DT	01-DEC-2001 (TREMblrel. 19, last annotation update)	
DE	PURATIVE ORGANIC ANION TRANSPORTER.	
GN	OAT1-4.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
OX	NCBI_Taxid=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=KIDNEY;	
RX	MEDLINE=20422308; PubMed=10964714;	
RA	Bain A., Prewitt D., Buttler D., Reid G., Enklaar T., Wolff N.A.,	
RA	Edlinghaus C., Hillemann A., Schulten H.J., Gunawan B., Fuezesi L.,	
RA	Zabel B., Burckhardt G.;	
RT	"Genomic structure and in vivo expression of the human organic anion	
RT	transporter 1 (hOAT1) gene.";	
RL	Biophys. Res. Commun. 275:623-630(2000).	
RD	EMBL: A0271205; CAB97245.1; -	
SO	SEQUENCE 519 AA; 57357 MW; E1748C6F9E2002P2 CRC64;	

[illegible]

RN [11] SEQUENCE FROM N.A.
 RA Rowen L., Madan A., Qin S., Baradarani L., Birditt B., Bloom S.,
 RA Burke J., Dors M., Fleetwood P., Kaur A., Madan A., Nesbitt R.,
 RA Pate D., Hood L.;
 Sequencing of human neurexin II gene.
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC044790; AAK68156.1; -
 SQ SEQUENCE 553 AA; 59630 MW; E3F29F38129BAD61 CRC64;

Query Match 42.7%; Score 601; DB 4; Length 553;

Best Local Similarity 37.8%; Pred. No. 1e-49; 81; Indels 108; Gaps 1;
 Matches 133; Conservative 30; Mismatches 81; Indels 108; Gaps 1;

QY 1 MAFELLISQVGLGRFOMLHVLPISMLLPHILLENFAAIPGRCVHMLDNTGS 60
 DB 1 MAFSELDLVGLGRFQVLOTALMVSIWMLCTQSMLENFAAIPGRCVAPLIDNSTAQ 60
 QY 61 GNETGISLSDALLRISIPDSNLREPKCRFFVHPQWQLHNTGISTSEADTEPCVDGW 120
 DB 61 ASILGSLSPALLAISIPGPNQRCRRFPQWQLDPPNATATSWSEADTEPCVDGW 120
 QY 121 VYDQSYFSTIYTKMDLVCDYQSLKSVQFILLTGMVGGIIGHSVDR----- 169
 DB 121 YDNRISFTSTIAKKNLVCDSHAKPKMAOSIYLAGILVGAAGCPADDRGRRLVLTWSY 180
 QY 170 ----- 169
 DB 181 LQMAVGTAAAFAPAPVYCLFREFLAFAVAGVMMNTGTLMEKTAARAPLVTLNLSLG 240
 QY 170 ----- 192
 DB 241 ESEGHGLTAAYAGVDWTLQLQVSPFFLCFLYSWMLAESAWMLTGTGLDMLGLOLW 300
 QY 193 KVARINGIKNAEETLNEVVRSTMOEELDAQTKTYCDLFRNSMKRRCIT 244
 DB 301 RVAALNGKAVQDTLPEVLLSAMREELSMGQPPASGLTLLRMPLGRLRPTCI 352
 RESULT 5
 Q9NSAO PRELIMINARY; PRT; 550 AA.
 AC Q9NSAO; 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE ORGANIC ANION TRANSPORTER 4 (OAT4).
 GN HOAT4 OR OAT4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1] SEQUENCE FROM N.A.
 RP TISSUE=KIDNEY;
 RC MEDLINE=20127945; PubMed=10660625;
 RA Cha S.H., Sekine T., Kusuhara H., Yu E., Kim J.Y., Kim D.K.,
 RA Sugiyama Y., Kanai Y., Endou H.;
 RT "Molecular cloning and characterization of multispecific organic anion
 transporter 4 expressed in the placenta.";
 RT J. Biol. Chem. 273:4507-4512(2000).
 RL [2]
 RN [2] SEQUENCE FROM N.A.
 RP Rowen L., Madan A., Qin S., Baradarani L., Birditt B., Bloom S.,
 RA Burke J., Dors M., Fleetwood P., Kaur A., Madan A., Nesbitt R.,
 RA Pate D., Hood L.;
 RT "Sequencing of human neurexin II gene.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC044790; AAK68155.1; -
 SQ SEQUENCE 550 AA; 59971 MW; 233BE6CA4520E58A CRC64;

Query Match 42.1%; Score 592; DB 4; Length 550;
 Best Local Similarity 40.2%; Pred. No. 7.7e-49;
 Matches 143; Conservative 28; Mismatches 69; Indels 116; Gaps 7;

QY 1 MAFELLISQVGLGRFOMLHVLPISMLLPHILLENFAAIPGRCVHMLDNTGS 58
 DB 1 MAFSELDLVGLGRFQVLOTALMVSIWMLCTQSMLENFAAIPGRCVAPLIDNSTAQ 56
 QY 59 GSGNETGISLSDALLRISIPDSNLREPKCRFFVHPQWQLHNTGISTSEADTEPCVD 118
 DB 59 GSAVSTN-KTPKALLTISIPGPNQRCRRFPQWQLDPPNATATSWSEADTEPCVD 115
 QY 119 GWYDQSYFSTIYTKMDLVCDYQSLKSVQFILLTGMVGGIIGHSVDR----- 169
 DB 116 GWYDQSYFSTIYTKMDLVCDYQSLKSVQFILLTGMVGGIIGHSVDR----- 175
 QY 170 ----- 169
 DB 176 CCLDLAAGTSTIPAPTVIYCGLRFVAFAGMAGIFLSITLMEVTTTSRAVTVTVG 235
 QY 170 ----- 190
 DB 236 CAFSAQAALGLAFALPDMRTLOLAASVFPFALISLWLPESARWMLIKGKPDALOE 295
 QY 191 LKVARINGIKNAEETLNEVVRSTMOEELDAQTKTYCDLFRNSMKRRCIT 246
 DB 296 LKVARINGIKNAEETLNEVVRSTMOEELDAQTKTYCDLFRNSMKRRCIT 350
 RESULT 6
 O54778 PRELIMINARY; PRT; 553 AA.
 AC O54778; 01-JUN-1998 (Tremblrel. 06, Created)
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE RST.
 GN SLC22A2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1] SEQUENCE FROM N.A.
 RP TISSUE=KIDNEY;
 RC MEDLINE=98072412; PubMed=9409754;
 RA Mori K., Ogawa Y., Ebihara K., Aoki T., Tamura N., Sugawara A.,
 RA Kuwahara T., Ozaki S., Mukoyama M., Tashiro K., Tanaka I., Nakao K.;
 RT "Kidney-specific expression of a novel mouse organic cation
 transporter-like protein.";
 RL FEBS Lett. 417:371-374(1997).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: AB005451; BAA23875.1; -
 DR MGD: MGI:1195269; SLC22A2.
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Transmembrane.
 SQ SEQUENCE 553 AA; 60160 MW; D991F5F7A439BBBF CRC64;

Query Match 39.2%; Score 551; DB 11; Length 553;
 Best Local Similarity 34.1%; Pred. No. 7.3e-45;
 Matches 120; Conservative 38; Mismatches 86; Indels 108; Gaps 1;

QY 1 MAFELLISQVGLGRFOMLHVLPISMLLPHILLENFAAIPGRCVHMLDNTGS 60
 DB 1 MAFSELDLVGLGRFQVLOTALMVSIWMLCTQSMLENFAAIPGRCVAPLIDNSTAQ 60
 QY 61 GNETGISLSDALLRISIPDSNLREPKCRFFVHPQWQLHNTGISTSEADTEPCVDGW 120
 DB 61 ASIPGLDGLPDLAVSIPGPNQRCRRFPQWQLDPPNATATSWSEADTEPCVDGW 120

[illegible]

PROTEIN	2		
096LX3			
ID	096LX3	PRELIMINARY;	PRT; 552 AA.
AC	096LX3;		
DT	01-DEC-2001	(TrEMBLrel. 19, Created)	
DT	01-DEC-2001	(TrEMBLrel. 19, Last sequence update)	
DT	01-DEC-2001	(TrEMBLrel. 19, Last annotation update)	
DE	CNDA FLJ33092.1FS, CLONE TRACH2000655, MODERATELY SIMILAR TO RATTS		
DE	NORVEGICUS PUTATIVE INTEGRAL MEMBRANE TRANSPORT PROTEIN (USTIR).		
OS	Homo sapiens (human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=TRACHEA;		
RA	Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohata N.,		
RA	Matsumoto K., Hirano M., Sano S., Momiyama R., Yoshikawa Y.,		
RA	Ra Matsunori Y., Moriya S., Chiba E., Nomiyama H., Onogawa S.,		
RA	Keiriyama S., Satoh N., Matsunawa H., Takahashi E., Kataoka R.,		
RA	Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,		
RA	Matsunabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Ota T.,		
RA	Mekamatsu A., Ishii S., Yamamoto J., Isono Y., Kawai-Hito Y., Saito K.,		
RA	Nishikawa T., Kimura K., Yamashita H., Matsuo K., Nakamura Y.,		
RA	Sekine M., Kikuchi H., Kanda K., Nagatsuma M., Murakawa K.,		
RA	Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,		
RA	Suzuki Y., Sugano S., Nagahara K., Masuno Y., Nagai K., Isogai T.,		
RT	"MEDO human cDNA sequencing project."		
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.		
SM	EMBL: AK075654; BAB71543.1; --		
SO	SEQUENCE 552 AA; 62140 MW; 39A831C0FD26CA68 CRC64;		

[illegible]

QY 169 -----RMVESAWLITTKDDEGKALR 192
Dd 241 ASIGMYIAGLAFPRIRHRLQLMSPYIPFELLTRMSESARWLVTYKPKQGLKELR 300
QY 193 KVARNGIKNAEELTIEVYASTMOEDLAQCTTYGOLFPRNPSMKRRCITVEFLR 249
Dd 301 KVALHNMGNKSGNTLTMEVVEASAKNELEAKRRSSPRDLPHFTLEKRRICVJLSFMR 357

RESULT	3	
091WJ2		
ID	091WJ2	PRELIMINARY;
AC	091WJ2	PRT; 552 AA.
DT	01-DEC-2001 (TREMBLrel. 19, Created)	
DT	01-DEC-2001 (TREMBLrel. 19, last annotation update)	
DT	01-DEC-2001 (TREMBLrel. 19, last annotation update)	
DE	SIMILAR TO SOLUTE CARRIER FAMILY 22 (ORGANIC CATION TRANSPORTER)-LIKE 2.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX	NCBI_TaxId:10090;	
NP	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE-KIDNEY;	
RA	Strausberg R.;	
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.	
SR	EMBL: BC014805; A0114805.1; .	
SO	SEQUENCE 552 AA; 61360 MW; F5D2498FC673EBFC CRC64;	

Query Match	Similarity	53.4%	Score 751	DB 11	Length 552
Best Local	Similarity 41.5%				
Matches 148	Conservative 44			Mismatches 57	Indels 108
					Gaps 1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 11:39:10 ; Search time 26.48 Seconds
(without alignments)
1750.855 Million cell updates/sec

Title: US-09-674-235-1

Perfect score: 1406
Sequence: 1 MAFFELLISVGLGRFQMLH.....RKIKSRKHKNDCTKTKTF 268

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL-19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	818	58.2	552	11 070609	070609 ratu
2	773	55.0	552	4 0961X3	0961X3 homo sapien
3	751	53.4	552	11 091WJ2	091WJ2 mus musculu
4	601	42.7	553	4 096S37	096S37 homo sapien
5	592	42.1	550	4 09NSA0	09NSA0 homo sapien
6	551	39.2	553	11 054778	054778 mus musculu
7	458.5	32.6	506	4 09NQC2	09NQC2 homo sapien
8	458.5	32.6	519	4 09NQA6	09NQA6 homo sapien
9	458.5	32.6	563	4 095742	095742 homo sapien
10	457.5	32.5	551	11 035956	035956 ratu
11	449.5	32.0	551	6 09TSY7	09TSY7 oryctolagus
12	439	31.2	562	13 057379	057379 pseudop
13	438.5	31.2	545	11 061185	061185 mus musculu
14	390.5	27.8	542	4 096TCL	096TCL homo sapien
15	372.5	26.5	552	4 096DT2	096DT2 homo sapien
16	368.5	26.2	537	11 091WJ9	091WJ9 mus musculu

17	367.5	26.1	537	11 088909	088909 mus musculu
18	362.5	25.8	536	11 09R1U7	09R1U7 ratu
19	324.5	23.1	535	11 063314	063314 ratu
20	318	22.6	540	11 091WU2	091WU2 mus musculu
21	312.5	22.2	140	11 09QYX0	09QYX0 ratu
22	307.5	21.9	539	4 09H2M5	09H2M5 homo sapien
23	306	21.8	548	4 09S820	09S820 homo sapien
24	296.5	21.1	548	4 09Y694	09Y694 homo sapien
25	234	16.6	551	4 09Y226	09Y226 homo sapien
26	228	16.2	538	5 09V6H5	09V6H5 drosophila
27	228	16.2	604	5 0961J5	0961J5 drosophila
28	223	15.9	548	5 001384	001384 drosophila
29	223	15.9	348	5 09VCA2	09VCA2 drosophila
30	220	15.6	564	11 09W7N6	09W7N6 mus musculu
31	202.5	14.4	553	11 09R1A1	09R1A1 ratu
32	197.5	14.0	553	11 09Z306	09Z306 mus musculu
33	197.5	14.0	568	5 09U539	09U539 caenorhabd
34	197.5	14.0	576	5 002270	002270 caenorhabd
35	192.5	13.7	567	5 09S8A8	09S8A8 drosophila
36	192	13.7	370	5 001383	001383 drosophila
37	191.5	13.6	567	5 09VCA3	09VCA3 drosophila
38	190.5	13.5	555	5 0961R9	0961R9 drosophila
39	190.5	13.5	557	4 096EH6	096EH6 homo sapien
40	189.5	13.5	577	4 096RU0	096RU0 homo sapien
41	184	13.1	177	4 096ER0	096ER0 homo sapien
42	182.5	13.0	502	5 09V8Z3	09V8Z3 drosophila
43	179.5	12.8	551	4 09H015	09H015 homo sapien
44	179.5	12.8	557	5 09VEX8	09VEX8 drosophila
45	178.5	12.7	551	4 014546	014546 homo sapien

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	552 AA.
070609	070609			
AC	01-AUG-1998 (TREMBLrel. 07, Created)			
DT	01-AUG-1998 (TREMBLrel. 07, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	PUTATIVE INTEGRAL MEMBRANE TRANSPORT PROTEIN.			
GN	USP1R.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=KIDNEY;			
RX	MEDLINE=98200080; PubMed=9541011;			
RA	Schoenly E., Spitzberger F., Engelhardt M., Martel F., Oeding N.,			
RA	Gruendemann D.,			
RT	"Molecular cloning and characterization of two novel transport			
RT	proteins from rat kidney."			
RL	FEBS Lett. 425:79-86(1998).			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.			
DR	EMBL: Y09945; CAI71076.1; -			
DR	InterPro: IPR003662; sub_transporter.			
DR	Pfam: PF00083; sugar_tr 1.			
KW	Transmembrane.			
SO	SEQUENCE 552 AA; 61476 MW; F47FBA2B41F5C4AF CRC64;			

Query Match 58.2%; Score 818; DB 11; Length 552;
Best Local Similarity 45.9%; Pred. No. 9.5e-71;
Matches 164; Conservative 40; Mismatches 45; Indels 108; Gaps 1;

QY 1 MAFFELLISVGLGRFQMLHVLPIISLMLIPHILENFAALIPGRVHMLDNTGS 60
Db 1 MAFFOLLNVOGSLGRFQMLHVLPIISLMLIPHILENFAALIPGRVHMLDNTGS 60